

STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 110061

TO: Georgia L Helmer
Location: 9d14 / 9e12
Wednesday, December 17, 2003
Art Unit: 1638
Phone: 308-7023
Serial Number: 09 / 805550

From: Jan Delaval
Location: Biotech-Chem Library
CM1-1E07
Phone: 308-4498

jan.delaval@uspto.gov

Search Notes

*OK
pending removed
12/17/03 glh*

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 06:16:40 ; Search time 18.3376 Seconds
(without alignments)
2123.954 Million cell updates/sec

Title: US-09-805-550-2
Perfect score: 2036
Sequence: 1 MCLNVKTLKGTNFEIASEPD.....BELTANYLLDHGHFDDQQQ 405

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1187.5	58.3	382	2 T14336	RAD23 protein, iso
2	1127	55.4	392	2 T04150	RAD23 protein homo
3	925.5	45.5	379	2 T14337	RAD23 protein, iso
4	840.5	41.3	367	2 F96827	protein F20B17.8 [
5	632	31.0	409	2 S44346	RAD23 protein homo
6	623	30.6	385	2 JC7783	RAD 23B protein -
7	588	28.9	368	2 T40115	uv excision repair
8	573	28.1	363	2 S44443	RAD23 protein homo
9	543	26.7	246	2 G86296	T24D18.27 protein
10	483.5	23.7	398	2 S50507	excision repair pr
11	436.5	21.4	372	2 T27774	hypothetical prote
12	361.5	17.8	113	2 H86296	F309.1 protein - A
13	229.5	11.3	551	2 C84549	probable ubiquitin
14	211	10.4	536	2 B84549	probable ubiquitin
15	202.5	9.9	502	2 T18562	hypothetical prote
16	193	9.5	142	2 T51479	hypothetical prote
17	168	8.3	373	2 S54583	ubiquitin-like pro
18	154.5	7.6	354	2 T38404	yeast dsk2 homolog
19	154.5	7.6	507	2 T44768	antifreeze glycope
20	150.5	7.4	581	2 T22341	hypothetical prote
21	149.5	7.3	1040	2 T29092	TSC-22 protein hom
22	147.5	7.2	424	2 T33663	hypothetical prote
23	147	7.2	292	2 T18584	F15C11.2 - Caenorh
24	146	7.2	681	2 A83455	DNA polymerase sub
25	144.5	7.1	852	2 T46091	hypothetical prote
26	144.5	7.1	907	2 AD2931	cell division prot
27	144.5	7.1	910	2 H98331	cell division prot
28	144	7.1	638	1 XXAV	dihydroliipoamide S
29	141.5	6.9	833	2 AP2089	hypothetical prote

30 141.5 6.9 2082 2 T37056
31 141 347 2 H75253
32 141 590 1 T35297
33 141 647 2 T39141
34 140.5 6.9 1116 2 S77213
35 140 6.9 638 2 H82690
36 140 6.9 2440 2 S39162
37 140 6.9 2441 2 S39161
38 139.5 6.9 1039 2 A12884
39 139.5 6.9 1132 2 A35098
40 139 6.8 340 2 A35630
41 138.5 6.8 762 2 H87302
42 138 6.8 421 2 A60058
43 137.5 6.8 547 2 H83018
44 137.5 6.8 865 2 A47282
45 137.5 6.8 873 2 A47283

ALIGNMENTS

RESULT 1

RAD23 protein, isoform I - carrot
T14336
C/Species: Daucus carota (carrot)
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C/Accession: T14336
R/Sturm, A.; Leinhardt, S.
Plant J. 13, 815-821, 1998
A/Title: Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in yeast.
A/Reference number: Z17989; MUID:98345997; PMID:9681019
A/Accession: T14336
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-382 <STU>
A/Cross-references: EMBL:Y12013; NID:g1914682; PIDN:CAA72741.1; PID:g1914683
A/Experimental source: subspecies Queen Anne's Lace, isolate W001C
C/Genetics:
A/Gene: RAD23-1
C/Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

Query Match 58.3%; Score 1187.5; DB 2; Length 382;
Best Local Similarity 62.7%; Pred. No. 2.9e-64;
Matches 255; Conservative 46; Mismatches 75; Indels 31; Gaps 11;
Qy 1 MCLNVKTLKGTNFEIASEPDASVADVKRIITETGQSTYRADQOQLYQSKILDETTLE 60
Db 1 MKIYVTKLGSQFEIQVNPDDSVADVKRSIETAQGAAYPAQQOQLYQSKVLKDGTTLL 60
Qy 61 SNGVAENSLVIMLSKAKASSSGASTATTA---KAPATLAQPAAPVAPASVARTPTQAP 117
Db 61 ENNVAENSLVIMLSKSKSPSGSTSTATAAPKAPOTSAPPSPV-APAVS-----QPP 113
Qy 118 VATAET-AP-PSVQPOQAPAAATVAATDDADVYSOASNLVFGNNLEQITQOILDMGGGTW 175
Db 114 ASTLPVPAPSPAPATAPIPSAAVGSEANVDSAAALLVAGSNLECAIQOILDMGGGTW 173
Qy 176 ERTVTVRAAAYNNPERAIDYLSGIPENVEQPVAPAPAGQQTWQOASPA-QPAPA 234
Db 174 DRDTVIRIVRAAFNNPERAVEYLSGIPQEAQAPPVAPSPSG-----QAANPLDQPPAA 228
Qy 235 LPQVPSASAGPNANPLNLPQGVPSGSGNPGVVPAGSGALDALROLPOFOALLQVQA 294
Db 229 --AQPAPASAGPNANPLDLFPQGLPDMGSN-----AAGAGNLDFLRNQQFOALRAVQVS 281
Qy 295 NPQILQMLQELGKQNPQILRLIQENQAEFLRLVNSPEGGPGGNILGQAAAAVPTLTIV 354
Db 282 NPQILQMLQELGKQNPQILRLIQEQHQAQDFLQINPEMGE--NLLGH-----GPOAISV 335
Qy 355 TPBERAIOBLEGNGFNRELVLEVFACNKKDEBELTANYLLDHGHEFD 401
Db 336 TPBERAIOBLEGNGFNRELVLEVFACNKKDEBELTANYLLDHGHEFE 382

pending
submitted
8/21/03

Query Match	45.5%;	Score 925.5;	DB 2;	Length 379;
Best Local Similarity	50.2%;	Pred. No. 1.6e-48;		
Matches	203;	Conservative	57;	Mismatches 117; Indels 27; Gaps 8;

Qy	1	MKLVNVLKGTNFEIEASPDASVADVKRIIETTTQGSTYRADDOOMLIYQGIKDKETTTLE	60
Db	1	MKLVNVLKGSHEIRVLPSPDTIMAVKKNIEDSGDKDNYPCGQQLLIHNGKVLKDESLA	60
Qy	61	SNQVAENSLFLVIMLSKAKA - SSSGASATTTAKAPATLAQPAAPVAPAAASVARTPTQAPVA	119
Db	61	ESKISEDGLFLVIMLGSKTSMTSGTTPAAQSSSAPATTPAPAVAPAPAPAAASAVIPNT	120
Qy	120	TASTAPSPVOQAPAAATVAATDDADYVSOAANLVFGNNLEOTIQOILDMGGGTWERDT	179
Db	121	TVPEAP - LSPAPAP - - - - -SDTYGSAASNVAGSLNEOTIQHIMDMGGGMDTNM	169
Qy	180	VPRALAAAYNNPRAIDLYLIGIPENVEAQ - PVARAPAAQOQTNQQAASPAQPAVALPVQ	238
Db	170	VSRALAAAYNNPRAVDLYLIGIPENAAAVPVSH - - FQDQINAGNNALSDNGVA - - -	223
Qy	239	PSPASAGPNANPLNPPQGVPGSGSNPGVVPAGSGALDALRLOPQALLQLLVQANPQI	298
Db	224	-GAAPCAPNSLPLNMPFQETLSG - - - - -VTGAGLSLEFLRNNPOFOTLRSVMVQRNPQI	276
Qy	299	LQPMLELKGKONPQIIRLLOENQAEFLRLVNESPEGPGGNIIGOLAAAVPOTLTVTPEE	358
Db	277	LQPMLELKGKONPQIIRLOQEHHEBFQLNLNPEVASE - GDMFDQPEDVQPSBITVTVAAD	335
Qy	359	REAIQRLGFMGRNRLVLEVFFACNKDDELTANYLLDHGHEPDD	402
Db	336	QEAIERLEAMGFDRGLVIEAFLACDRNEELAVNYLLENAGDFED	379

RESULT 4	
F96827	protein F20B17.8 [imported] - Arabidopsis thaliana
C:Species:	Arabidopsis thaliana (mouse-ear cress)
C:Date:	02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002
C:Accession:	F96827
R:Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; D. ansen, N.F.; Hughes, B.; Huizar, L.	
Nature	408, 816-820, 2000
A:Authors:	Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors:	Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title:	Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number:	A86141; MUID:21016719; PMID:11130712
A:Accession:	F96827
A>Status:	preliminary
A:Molecule type:	DNA
A:Residues:	1-367 <STO>
A:Cross-references:	GB:AE005173; NID:g7715605; PIDN:AAF68123.1; GSPDB:GN00141
C:Genetics:	
A:Gene:	F20B17.8
A:Map position:	1
C:Superfamily:	unassigned ubiquitin-related proteins; ubiquitin homology

Query Match	41.3%;	Score 840.5;	DB 2;	Length 367;
Best Local Similarity	47.4%;	Pred. No. 1.9e-43;		
Matches	192;	Conservative	53;	Mismatches 119; Indels 41; Gaps 8;

Qy	1	MKLVNVLKGTNFEIEASPDASVADVKRIIETTTQGSTYRADDOOMLIYQGIKDKETTTLE	60
Db	1	MKLVNVLKGSHEIRVLPSPDTIMAVKKNIEDSGDKDNYPCGQQLLIHNGKVLKDESLV	60
Qy	61	SNQVAENSLFLVIMLSKAKASSSGASATTTAKAPATLAQPAAPVAPAAASVARTPTQAPVA	120
Db	61	ENKVTBEGFLVIMLSKSKSGGS - - - - -AQASQVCVRLLLFHSFLPFLHRLSSINYPVFT	116

QY 121 AETAPSVQPAAPATVAATDDADYVSOAASNLVFGNNLEQTIQOILDMGGTWERDVT 180
Db 117 CSCFTYS-----CSTTGTDTYGAASTLVSGSSLEQMVQOIMEGGSGMDKETV 166
QY 181 VRALRAAYNNPERAIDYLYSGIPENVE-AQPVARAPAAGQQTNQQAASPAQPAVALPVQP 239
Db 167 TRALRAAYNNPERAIDYLYSGIPQTAEVAVPPEAQIAG-----SGAAPV-- 211
QY 240 SPASAGNANPLNLFQGVPSGSGNPGVVGAGSGALDALRQI-POFQALLQVLQVQANPOI 298
Db 212 APASGGPNSPLDLFFQ-----ETVAAGSGDLGTLFLRNNDQVAILTISAFSLN 262
QY 299 LQPMLELQKQNPQIIRLQENQAEFLRLVNESPEGGPG-GNLLGQLAAAVPOTLTVTPE 357
Db 263 CEPMLQELQKQNPQLRLQENQAEFLRLVNEPEYEGSDGDMFQDQPEMPHAINVTFA 322
QY 358 ERAIORLEQMGFNRLVLEVPFACNKBELTANYLLDHGHEPDD 402
Db 323 EQEAIORLEAMGFDRALVIEAFIACDRNEELANYLLENSGDPED 367
RESULT 5
S44346
RAD23 protein homolog - human
N/Alternate names: p58/HHR23B; XP-C repair complementing protein
C/Species: Homo sapiens (man)
C/Date: 25-Dec-1994 #sequence_revision 01-Sep-1995 #text_change 24-Sep-1999
C/Accession: S44346
R/Masutani, C.; Sugawara, K.; Yanagisawa, J.; Sonoyama, T.; Ui, M.; Enomoto, T.; Takio, EMO J. 13, 1831-1843, 1994
A/Title: Purification and cloning of a nucleotide excision repair complex involving the A/Reference number: S44345; MUID:94222030; PMID:8168482
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-409 <WAS>
A/Cross-references: EMBL:D21090; NID:g498147; PIDN:BAA04652.1; PID:d1005181; PID:g498148
C/Superfamily: unassigned ubiquitin-related proteins; ubiquitin homolog
F/1-80/Domain: ubiquitin homolog <UBH>
Query Match 31.0%; Score 632; DB 2; Length 409;
Best Local Similarity 34.8%; Pred. No. 7.4e-31; Indels 70; Gaps 11;
Matches 150; Conservative 82; Mismatches 138;
QY 1 MKNLVTKLTGNTFEIASPDASVADVVKRIETTQGSQTYRADQOQLIYQKILKDETTLE 60
Db 1 MQVTLTQQOQTKIDIDPEETVKALKKEIESEKDAFPVACQKLIYAGKILNDDTALK 60
QY 61 SNGVAENSFLVIMLSKAKASSSGASTATTAKAPA-----TLAQAAPVAPAAV 109
Db 61 EYKIDKFNFVVMVTKPKAVSTPAPATTQGSAPASTTAVTSSTTTVAQAPTVP-PALAP 119
QY 110 ARTPTQ-----APVATAETAPSVQPAAPAT-VAATDD-----ADVYSQ 149
Db 120 TSPASITPASATSEPAASAKAQEPAETAVTATPTADTSDGDSRSNLFED 179
QY 150 AASNLVFGNNLEQTIQOILDMGGTWERDVTVRALRAAYNNPERAIDYLYSGIPENVEAQ 209
Db 180 ATSALTVGSGYENMVTEIMSG---YERQVVAALRASFPNDRAVEYLLTGIPVESEGNVGASDPAA 236
QY 210 PVARAPAAGQQTNQQAASPAQPAALVQVPSASAGNANPLNLFQGVPSGSGNPGVVP 269
Db 237 AVVDPPQAA-----STGAQSSAVAAAAATTATT-----TTTSSGGHP---- 275
QY 270 GAGSGALDALRQLPQFALLQVLQVQANPQILQELQKQNPQILRLIQENQAEFLRLVN 329
Db 276 -----LEFLRNQFQPMQRITIQNNPSSLIPALLOQIGRENPOLLOQISSHQEHFQMLN 329
QY 330 E-----SPGGGPGNLTQLAAVQ---QTLTVPTEERAIORLECGMGNRLVLEVFVA 381
Db 330 EPVQEAAGGGGGGGGGGGAAGSGHMNYIQVTPOKEAERLKAUGFPPEGLVIQAYFA 389
QY 382 CNKDELTANYLLDHGHEFD 401

Db 390 CEKNENLAANFLQQNFDE 409
RESULT 6
JC7783
RAD 23B protein - channel catfish
C/Species: Ictalurus punctatus (channel catfish)
C/Date: 02-Apr-2002 #sequence_revision 02-Apr-2002 #text_change 02-Apr-2002
C/Accession: JC7783
R/Liu, Z.; Li, P.; Kocabas, A.; Karsi, A.; Ju, Z.
Biochem. Biophys. Res. Commun. 289, 317-324, 2001
A/Title: Microsatellite-containing genes from the channel catfish brain: Evidence of tr A/Reference number: JC7783
A/Contents: Brain
A/Accession: JC7783
A/Molecule type: mRNA
A/Residues: 1-385 <LIU>
C/Comment: This protein with a polythreonine tract, has importance in the nucleotide ex C/Genetics:
A/Gene: rad23b
A/Introns: 76/3
Query Match 30.6%; Score 623; DB 2; Length 385;
Best Local Similarity 34.6%; Pred. No. 2.4e-30; Indels 64; Gaps 11;
Matches 147; Conservative 85; Mismatches 129;
QY 1 MKNLVTKLTGNTFEIASPDASVADVVKRIETTQGSQTYRADQOQLIYQKILKDETTLE 60
Db 1 MQVTLTQQOQTKIDIDAEETVKALKKEIENKESFPVAGQKLIYAGKILNDDTALK 60
QY 61 SNGVAENSFLVIMLSKAKASSSGASTATTAKAPATLAQPAAPVAP--AASVARTPTQAP 117
Db 61 EYKIDKFNFVVMVAKPKAATAAQAQSSSTTAASSSSSTSTTTTTPVPPVAASAATTTT 120
QY 118 VAT-----AETAPPVQPAAPATVAATDDADYVSOAASNLVFGNNLEQ 163
Db 121 TTTTSTSESVIEKAAEEKPPSTPASGSLT-----NVNIFEATSALTVGSGYEMM 175
QY 164 IQOILDMGGTWERDVTVRALRAAYNNPERAIDYLYSGIPENVEAQPAAPAAQQTNQ 223
Db 176 VTEIMLMG---YERQVVAALRASFPNDRAVEYLLTGIPVESEGNVGASDPAA----- 226
QY 224 QAASPAQPAVALPVQVPSASAGNANPLNLFQGVPSGSGNPGVVPVFGAGSGA--LDALRQ 281
Db 227 -----PVGAPA-----VTTGLSSPSTTPTQTAGSGANPLEFLRN 263
QY 282 LPFOALLQVLQVQANPQILQELQKQNPQILRLIQENQAEFLRLVN-----ESPEGGPG 337
Db 264 QPQFLQMRQITIQNNPSSLIPALLOQIGRENPOLLOQISSHQEQFIQMLNEPAQEAQGGGG 323
QY 338 GNILGQLAAAVPQTLTVPTEERAIORLECGMGNRLVLEVPFACNKBELTANYLLDHG 397
Db 324 G--VSEAGGGHMNYIQVTPOKEAERLKAUGFPPEGLVIQAYFACEKNLAANFLQ-- 379
QY 398 HEFDD 402
Db 380 QNFDD 384
RESULT 7
T40115
uv excision repair protein rad23 homolog - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Sep-2000
C/Accession: T40115; T51298
R/Wood, V.; Rajandream, M.A.; Barrell, B.G.; Taylor, K.; Harris, D.
submitted to the EMBL Data Library, September 1998
A/Reference number: Z21906
A/Accession: T40115
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-368 <WOO>

Qy	1	MKLNVTKLKGTNEIEASPDASVADVVKRIIETTGQSGTYRADQOMLIYQGKILKDETTLE	60
Db	1	MKLNVTKLKGTNEIEASPDASVADVVKRIIETTGQSGTYRADQOMLIYQGKILKDETTLV	60
Qy	61	SNQVAENSLVIMLSKAKASSSGASTATTAKAPATLAQPAAPVAPAAASVARTPTQAPVAT	120
Db	61	ENKVTBEGFLVWLSKSKTASSAGPSTQLAASTTQSIAVPASNSTPVOEQPTA	115
Qy	121	AETAPPSVQQAAPAAATVADDDADVTYSQAASNLVFGNNLEQTIQIILDMGGGTWERTDV	180
Db	116	-----QSDTYGQAASTLVSGSSIEQWVQVQIMEMGGSGGSDKQETV	153
Qy	181	VRALRAAYNNPERAIDLYLSGIPENVEAQPVARAPAAGQQTNOQAASPAQPAVALPVQPS	240
Db	154	TRALRAAYNNPERAIDLYLSGIPETV---TIPATNLGSGVSGRELTAP-----	198
Qy	241	PASAGPNANPLNLFPOQVPSGGSNPGVVPVPGAGSGALDALR	280
Db	199	PPSGGPNSSPLDLPFQPAVSDAA-----GGDLGTLEFLR	232

RESULT 10
 S50507
 excision repair protein RAD23 - yeast (Saccharomyces cerevisiae)
 N;Alternate names: protein YEL037C
 C;Species: Saccharomyces cerevisiae
 C;Date: 28-Jan-1995 #sequence revision 12-May-1995 #text_change 24-Sep-1999
 C;Accession: S50507; S30845; A54599; S38547
 R;Dietrich, F.S.
 submitted to the EMBL Data Library, December 1994
 A;Description: The sequence of S. cerevisiae cosmid 8199, 8334, and 9871.
 A;Reference number: S50491
 A;Accession: S50507
 A;Molecule type: DNA
 A;Residues: 1-398 <DIE>
 A;Cross-references: EMBL:U18779; NID:g603625; PIDN:AAB65005.1; PID:g603642; MIPS:YEL037C
 R;Mulligan, J.T.; Dietrich, F.S.; Hennessey, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Taylor, R.; Watkins, J.F.; Sung, P.; Prakash, L.; Prakash, S.
 submitted to the EMBL Data Library, February 1993
 A;Reference number: S30812
 A;Accession: S30845
 A;Molecule type: DNA
 A;Residues: 1-398 <MUL>
 A;Cross-references: GB:U18779; EMBL:LI0830; NID:g603625; PIDN:AAB65005.1; PID:g603642
 R;Watkins, J.F.; Sung, P.; Prakash, L.; Prakash, S.
 Mol. Cell. Biol. 13, 7757-7765, 1993
 A;Title: The Saccharomyces cerevisiae DNA repair gene RAD23 encodes a nuclear protein
 A;Reference number: A54599; MUID:94067136; PMID:8246991
 A;Accession: A54599
 A;Molecule type: DNA
 A;Residues: 1-398 <WAT>
 A;Cross-references: GB:L25428; NID:g409246; PIDN:AAA16070.1; PID:g409247
 R;Welnick, L.; Sherman, F.
 J. Mol. Biol. 233, 372-388, 1993
 A;Title: The gene clusters ARC and COR on chromosomes 5 and 10, respectively, of Saccharomyces cerevisiae
 A;Reference number: S38543; MUID:94016558; PMID:8411151
 A;Accession: S38547
 A;Molecule type: DNA
 A;Residues: 1-276 'R', 278-398 <MEL>
 A;Cross-references: EMBL:S66117; NID:g430823; PIDN:AAB28441.1; PID:g430824
 C;Genetics:
 A;Gene: SGD:RAD23
 A;Cross-references: SGD:S0000763; MIPS:YEL037C
 A;Map position: 5L
 C;Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology
 C;Keywords: DNA repair
 F;2-77/Domain: ubiquitin homology <UBH1>

Query Match	23.7%	Score	483.5;	DB 2;	Length	398;			
Best Local Similarity	32.3%	Pred.	No. 5.9e-22;						
Matches	142;	Conservative	65;	Mismatches	144;	Indels	89;	Gaps	18;

Qy 1 MKLNVTKLKGTNEIEASPDASVADVVKRIIETTGQSGTYRADQOMLIYQGKILKDETTLL 59

Db 2 VSLTFKNFKKXVPLDLPEPNTILETK-----TKLAQSI SCBESQIKLYSGKVLQDSKTV 57

Qy 60 ESNQVAENSFLVIMLSKAKASSGASTATTAKAPATLAQPAAPVAPAAASVARTP-----113

Db 58 SECGLKGDQVFWVWSQK-----STKTKVTEP-----PIAPES--ATTPGRENST 101

Qy 114 -----TQAPVATARTAPSPVQPOAPAAATVAATDADVYVSOAASNLVFGNNLEQTIQIL 168

Db 102 EASPSTDASAAPATAPEGSQPEQATATTERTESA-----STPGFVVYGTERTETIERIM 156

Qy 169 DMGGTWERDVTVRALAAAYNNPERAIDLYLSGIPENVEAQPVARAPAAAGQOTNOQAASP 228

Db 157 EMG---YQREVERALRAAFNNPDRAVEYLLMGIPENL-----RQPEPQQT-----AAMA 204

Qy 229 AQPAVALPVQFSPASAGFNPANPLNLFQGVPSGGSNPGVPGAGSGALDALRQLP-----283

Db 205 EQPSTAATTAEQPAED-----DLFAQAAGQGNASSGAL-GTTGGATDAAQGGPPGSTG 256

Qy 284 -----QFQALLQLVQANFOILPMLOELGKQNPQILRLIQENQAEPLRLVNES-----331

Db 257 LTVBDLLSLRQVSGNPEALAPLLENISARYPOLREHIMANPEVFSVLMLEAVGDNMQDV 316

Qy 332 PEG-----GPGGNILGQAAA-VPO-----TLTVTPEREATORLECGMGNRELVL 377

Db 317 MEGADDMVEGEDIEVTGEAAAGUGGEGEGSFQVDYTPDDQQAISRUCELGFERDLVQ 376

Qy 378 VFFACNKDEELTANYLL-DH 396

Db 377 VYFACDKNEEAANILFSDH 396

RESULT 11

T27774

hypothetical protein ZK20.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000

C;Accession: T27774

R;Gajadaty, S.

submitted to the EMBL Data Library, March 1996

A;Reference number: Z20417

A;Accession: T27774

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-372 <WIL>

A;Cross-references: EMBL:Z69904; PIDN:CAA93780.1; GSPDB:GN00020; CESP:ZK20.3

A;Experimental source: clone ZK20

C;Genetics:

A;Gene: CESP:ZK20.3

A;Map position: 2

A;Introns: 49/2; 73/3; 101/3; 369/3

C;Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

Query Match 21.4%; Score 436.5; DB 2; Length 372;

Best Local Similarity 28.9%; Pred. No. 3.6e-19;

Matches 117; Conservative 71; Mismatches 120; Indels 97; Gaps 13

Qy 1 MKLNVKTLKGNFNFEIASPDAADVKKRIETTGQSTYRADQQMLIYQKILKDETLE 60

Db 52 LSVTFRLTQVNFNLNEDQTIAEVKALVASEKDD-YAPELQKLIYNGKILDDSVKVG 110

Qy 61 SNGVAENSFLVIMLSKAKASSGASTATTAKAPATLAQPAAPVAPAAASVARTPTQAPVAT 120

Db 111 EVGPDSSKFVWMLSKRKVTEVAPSSTV-----ATAAEP-VPVAAA-----PASNPAPA 158

Qy 121 AETAPSPVQPOAAAATVAATDADVYVSOAASNLVFGNNLEQTIQIILDMGGTGTERTV 180

Db 159 ADVAPEA-----AAPAEAFALDEQ-----BENVLAITGMG---YDREQT 195

Qy 181 VRALRAAYNNPERAIDLYLSGIPENVEAQPVARAPAAAGQOTNOQAASPAQPAVALPVQPS 240

Db 196 IALRALRAAFNPNDRAVEFLNLGIPDD-----AAQEPD-----227

Qy 241 PASAGPNANPNLNFQGVPSGGSNPGVVPAGSGALDALRQLPQFQALLQLVQANFQILQ 300

```
Db 228 ---LGPEQNIDN-----VDGNDLNNMLANNPQALAEIRALIQNPMLA 269
Qy 301 PMLQELGKQNPQLRLIQENQAEFLRLVNESPEGPGGNILGOLAAAAPQTLT----- 353
Db 270 AVLQQLAAVNPRLVQTIONNQAFMDLLNGAQQGA-----GAAAGNAPERNTPRRHVIH 323
Qy 354 VTPEREAIQRLGEGFN--RELVLVEFFACNKDELTANYLLDH 396
Db 324 LSPPEAAAIERIKAIIVNAPEAVVVEAYFACDKNEEAINFISN 368

RESULT 12
H86296
F309.1 protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
R:Accession: H86296
R:Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, C.J.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: H86296
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-113 <STO>
A:Cross-references: GB:AE005172; NID:94966345; PIDN:AAD34676.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 17.8%; Score 361.5; DB 2; Length 113;
Best Local Similarity 61.1%; Pred. No. 2.7e-15;
Matches 69; Conservative 22; Mismatches 21; Indels 1; Gaps 1;

Qy 291 LVQANPQLQPMQLGKQNPQLRLIQENQAEFLRLVNESPEGPGG-NILGQAAAAPV 349
Db 1 MVNSNPQLQPMQLGKQNPQLRLIQENQAEFLRLVNESPEGPGGVDIFDQDQEMP 60

Qy 350 QTLTTPPERAIQRLGEGFNRELVLVEFFACNKDELTANYLLDHGHEPDD 402
Db 61 HSNVTPPERQESIERLEANGFDRAIVIEAFUSCDRNEELAAANYLLEHSADFED 113

RESULT 13
C84549
probable ubiquitin-like protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
R:Accession: C84549
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: C84549
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-551 <STO>
A:Cross-references: GB:AE002093; NID:94584343; PIDN:AAD25138.1; GSPDB:GN00139
C:Genetics:
A:Map position: 2

Query Match 11.3%; Score 229.5; DB 2; Length 551;
Best Local Similarity 23.7%; Pred. No. 1.6e-06;
```

```
Matches 132; Conservative 61; Mismatches 170; Indels 193; Gaps 26;
Qy 3 LNVKTLKGTNFEIEASPDASVADVKRIITETGOSTYRADQOQMLYQGKILKDETTLESN 62
Db 20 VNRCSNGTKFSVTKSLDSTVESFKELVAQS---SDVPANQORLIYKGRILKDDQTLSSY 76
Qy 63 GVAENSLVIMLSKAKASSGASTATTAKAPATLQAAPAVAPASVARPTQAPVATAE 122
Db 77 GLQADH--TIHMVRGSAPSS-----APPAPASQTTAPSVTRGVGSD 117
Qy 123 TA-----PPSVQPOAAAPAAVTAATDDADVY 147
Db 118 NSSNLGGASPGESLFGGLGNPLGGNMSGLFGAGLPDLVQTQQQLAQNPNMIRDM-MN 176
Qy 148 SQAASNLVFGNNLE-----QTIQIILDMG---GGTWERDTVVRALRAAYNPE--- 192
Db 177 TPAQLNLM--NNPEFMSMIMNPNQMBELVDNRNPELGHVLDNPISILRQTLAEARNPELMR 234
Qy 193 ---RAIDYLSGIP-----ENVEAQVARAPAAQOQTNOQAAAPQAPAV--- 233
Db 235 EMERNTRAMSNIESMPEGFNMLRMYENVQEPMLNATTMSGNAGNNTGSPNPPFAALLGNQ 294
Qy 234 ALPVQSPAS-----AG-----PNANPLNLPQGVPSG-----GSPGV--- 267
Db 295 GVTTQGSADSNSSSTPNAGTGTIPNANPLP-NPWGATGGQTAPGRTNVVGSDARSPLGG 353
Qy 268 VPGAGS-----GA-----LDALRQLPQFOALLQLVOANPQI-----L 299
Db 354 LGGLSGLGGGLGMLGADSPGATPDASQSLQLQNPASQMMQSVFSNPNQYNQMLMSL 413
Qy 300 QPMQLGKQNPQLRLIQENQAEFLRLVNESPE----- 333
Db 414 NPQLRSLDSDNPQLREMMQ--NPDLRQFS--SPENMQMMTLOQSLSNRNTASQDAQOT 470
Qy 334 ---GGPGG-----NILGOLAA---AVPQTLTTPPEREA---IQRLGEGF-NRELVLVEV 378
Db 471 GAATGNNGLDLMNMFSGAGGLSGTNSQSVNPPERYATQLOQLQEMGFYDRAENIRA 530
Qy 379 FFACNKDELTANYLL 394
Db 531 LLATNGVNNAVERLL 546

RESULT 14
B84549
probable ubiquitin-like protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
R:Accession: B84549
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: B84549
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-536 <STO>
A:Cross-references: GB:AE002093; NID:94584342; PIDN:AAD25137.1; GSPDB:GN00139
C:Genetics:
A:Map position: 2

Query Match 10.4%; Score 211; DB 2; Length 536;
Best Local Similarity 21.5%; Pred. No. 2e-05;
Matches 125; Conservative 56; Mismatches 142; Indels 258; Gaps 23;
Qy 3 LNVKTLKGTNFEIEASPDASVADVKRIITETGOSTYRADQOQMLYQGKILKDETTLESN 62
Db 20 VNRCSNGTKFSVTKSLDSTVESFKELIAQ---NSDVPANQORLIYKGRILKDDQTLSSY 76
Qy 63 GVAENSLVIMLSKAKASSGASTATTAKAPATLQA-----PAAPVAPASVARTPTQAP 117
```

```

Qy 60 ESNQVAENS--FLVIMLSKAKASSGASTATTAKAPATLAQPAAPVAP- - - - - 106
Db 64 TQKIADGHTVHLVIRNQARPTPAPAAATPTASSAPSSNPFPSSQPNPTNPFPAAMGGMG 123
Qy 107 --ASVARTP-----TQAPVATAETAAPPSPVQPAAPAAATVAAT- - - - -DDADV 146
Db 124 SPADILNNPDMRSMVDNPITTOQLLGNPEFMRITIQSNPQFQALIERNPVEGHIILNDPNV 183
Qy 147 YSOASNLVFN-----NLEQITQIILDGGG--TWER---DTVVRLRAAYN--N 190
Db 184 MRQTMETIRNPMFQEMRNRHDQAIRNLQIPGGEAALERLYNDVQEPLLNSATNSLGN 243
Qy 191 PERAIDVLYSGIP-----ENVEAQPARAPAGAQQTNOA- - - - -ASPAQA 232
Db 244 PFASIRGQDSSEPRVDRAGQENNEALPNFNASNANQATNNSNNRSADFNSLLDSPGISS 303
Qy 233 VALPVQSPASAGNPANP--LNLFPQGVSPSGSNPGVVPGAGSGALDALRQPQ- - - - - 284
Db 304 LMEQWMSNPMSQASMFSEPVINSRQNM--SNPFGLDSI-VGQIPARDNPQISEGR 359
Qy 285 --FOALLQLVQANPQIILQPMQLGKQNPQIILRIQENQAQEPRLVNSPE-----GG 335
Db 360 RSFPQMLNM--SDPSVMEAM-----RNPVRSEAFROTQEGFSTLRREAPOLLNLFQAGA 412
Qy 336 PGGNIGQLAAAV-----POTLH--VTPEREH--IORLSEGM 368
Db 413 MGGGAFSGSDANASSAGANSANGLADLFNSMMNGGGRPSSTAAPVNPQETIYASQLEQLQSM 472
Qy 369 GFN 371
Db 473 GFS 475

```

Search completed: December 17, 2003, 06:20:48
Job time : 21.3376 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 06:16:39 ; Search time 11.5265 Seconds
(without alignments)
1652.347 Million cell updates/sec

Title: US-09-805-550-2

Perfect score: 2036

Sequence: 1 MKLVNKLKGTNFEIASPD.....EELTANYLLDHGHEFDQQQ 405

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	632	31.0	409	1 R23B_HUMAN	P54727 homo sapien
2	623.5	30.6	416	1 R23B_MOUSE	P54728 mus musculus
3	588	28.9	368	1 RH23_SCHPO	O74803 schizosacch
4	577	28.3	363	1 R23A_MOUSE	P54726 mus musculus
5	573	28.1	363	1 R23A_HUMAN	P54725 homo sapien
6	483.5	23.7	398	1 RA23_YEAST	P32628 saccharomyc
7	167	8.2	373	1 DSK2_YEAST	P48510 saccharomyc
8	154.5	7.6	354	1 YAUG_SCHPO	Q10169 schizosacch
9	149	7.3	1083	1 T2D3_HUMAN	O00268 homo sapien
10	144	7.1	637	1 ODP2_AZOV1	P10802 azotobacter
11	140.5	6.9	1509	1 GSRI_HUMAN	Q9NZM4 homo sapien
12	140	6.9	2441	1 CBP_MOUSE	P45481 mus musculus
13	139.5	6.9	1132	1 BAT3_HUMAN	P46379 homo sapien
14	137.5	6.8	547	1 ODP2_PSEAE	O59638 pseudomonas
15	137.5	6.8	552	1 EFG1_CANAL	P43064 candida alb
16	137.5	6.8	865	1 CPN_DROME	O02910 drosophila
17	137.5	6.8	1012	1 PHC1_MOUSE	Q64028 mus musculus
18	136	6.7	743	1 P021_HUMAN	P14859 homo sapien
19	136	6.7	797	1 PAT1_YEAST	P25644 saccharomyc
20	135	6.6	352	1 ALGP_PSEAE	P15276 pseudomonas
21	135	6.6	1048	1 SRA4_RAT	Q63627 rattus norv
22	134.5	6.6	699	1 VGLG_HSV2H	P13290 herpes simp
23	134	6.6	361	1 IFP3_MOUSE	O9DC44 mus musculus
24	133.5	6.6	1157	1 SRA4_HUMAN	O92104 homo sapien
25	133.5	6.6	2442	1 CBP_HUMAN	Q92793 homo sapien
26	131	6.4	1783	1 RAA3_CHLRE	Q9FEC4 chlamydomon
27	130	6.4	392	1 HME1_HUMAN	Q05925 homo sapien
28	129.5	6.4	1772	1 MSP1_PLAYO	P13828 plasmodium
29	129	6.3	511	1 P60_LISGR	Q01835 histaria gr
30	129	6.3	688	1 DP3X_HAEIN	P43746 haemophilus
31	129	6.3	1794	1 RDM5_SCHPO	Q10172 schizosacch
32	127.5	6.3	815	1 RBMS_HUMAN	P52756 homo sapien
33	127	6.2	497	1 HME5_DROME	P18488 drosophila

RESULT 1

R23B_HUMAN
ID R23B_HUMAN STANDARD; PRT; 409 AA.
AC P54727; Q8WUB0;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE UV excision repair protein RAD23 homolog B (HHR23B) (XP-C repair
DB complementing complex 58 kDa protein) (P58).
GN RAD23B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=94222030; PubMed=8168482;
RA Masutani C., Sugawara K., Yanagisawa J., Sonoyama T., Ui M.,
RA Konomoto T., Takio K., Tanaka K., van der Spek P.J., Bootsma D.,
RA Hoeijmakers J.H.J., Hanaoka F.;
RT "Purification and cloning of a nucleotide excision repair complex
RT involving the Xeroderma pigmentosum group C protein and a human
RT homologue of Yeast RAD23";
RL EMBL J. 13:1831-1843 (1994).
[2]
SEQUENCE FROM N.A., AND VARIANT VAL-249.
RP Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung M.-W.,
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
RA Schackwitz W.S., Sherwood J.K., Wirak L.A., Nickerson D.A.;
RN Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
RP Ramsay H.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE FROM N.A., AND VARIANT VAL-249.
RP TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length

ALIGNMENTS

34	127	6.2	632	1	P021_RAT
35	127	6.2	1544	1	TUSP_HUMAN
36	126.5	6.2	721	1	YK82_MYCTU
37	126.5	6.2	825	1	ICP0_HSV2H
38	126.5	6.2	1115	1	NCAL_MOUSE
39	126	6.2	788	1	PCAP_HUMAN
40	125.5	6.2	1211	1	BUN2_DROME
41	125.5	6.2	2090	1	N214_HUMAN
42	125	6.1	2459	1	MAPB_RAT
43	124.5	6.1	3149	1	TEGU_EBV
44	124	6.1	2414	1	P300_HUMAN
45	123.5	6.1	1596	1	MAM_DROME

P31503 rattus norv
Q9nrj4 homo sapien
Q10690 mycobacteri
P28284 herpes simp
P13595 mus musculu
Q96rn5 homo sapien
Q24523 drosophila
P35658 homo sapien
P15205 rattus norv
P03186 epstein-bar
Q09472 homo sapien
P21519 drosophila


```

DR PROSITE; PS50030; UBA; 2.
DR DNA damage; DNA repair; Nuclear protein; Repeat.
FT DOMAIN 1 79 UBIQUITIN-LIKE.
FT DOMAIN 188 228 UBA 1.
FT DOMAIN 371 411 UBA 2.
FT DOMAIN 255 261 POLY-ALA.
FT DOMAIN 282 270 POLY-THR.
FT DOMAIN 336 355 POLY-GLY.
SQ SEQUENCE 416 AA; 43516 MW; 1380245A6D892205 CRC64;

Query Match          30.6%; Score 623.5; DB 1; Length 416;
Best Local Similarity 32.4%; Pred. No. 2.5e-28;
Matches 146; Conservative 85; Mismatches 136; Indels 83; Gaps 10;

QY 1 MKNVTKLTGNTFEIEASPDASVADVXRIETQGSTVRADQMLIIQGIKDKDTTLE 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MQVTLKTLQQQTFKIDIDPEETVKALKEIESEKGDAPFVAGKLIYAGKILSDTALK 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 61 SNGVAENSFLVIMLSKAKA-----SSGSASTATTAKAPATLAQPAAPVAPASVA 110
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 EYKIDKKNFVVMVTKRAVTTAVPATTPQSPSTPSTTVSSPAV-----AAQAAPATPA 116
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 111 RTPQAPVAT-----AETAP-----PSVQPOQAPATVAATDD-----ADV 146
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 117 LAPSTPASTTPASTTASSEPAPAGATQPEKPAEQTPVLTSPADSTPGDSSRNL 176
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 147 YSQAASNLVFGNNLQTTQIILDMGGTWERDVTVRALRAAYNNPERAIDLYLGIPNV 206
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 177 FEDATSAVTGQSYENWTEINSMG---YERQVIAALRASFNPNDRAVEYLLMGIPDR 233
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 207 EAQPVARAPAGQQTNQQAASPAQPAVALPVQSPASAGPNANLPLFPQGVPSGSGNP 266
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 234 ESQAVVDPQPAVST---GTPQSPAVAAAAATTTATT-----TTTSGGHP- 275
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 267 VVPGAGSALDALROLPOFQALLQVQNPQILQPMQLGKQNPQIIRLQENQAEFLR 326
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 276 -----LEPLRNQPOFQNRQIITQNPSPALLPALQOIGRENQQLAQISQHQEHP 326
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 327 LVNESPE-----GGPGNLTGLQAAAVQTTLVTPEEREAIORLEGMGFN 371
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 327 MLNEVPQAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 371
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 372 RELVLEVFACNKBELTANYLLLDHGRFD 401
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 387 EGLVITQVAFACEKNENLAANFLQQNFED 416
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 3
RH23 SCHPO
ID RH23 SCHPO STANDARD; PRT; 368 AA.
AC O74803;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE UV excision repair protein rhp23 (RAD23 homolog).
GN RHP23 OR SPBC2D10.12.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]_
RP SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
RX STRAIN=SP223;
RC MEDLINE=21648918; PubMed=11788722;
RA Elder R.T., Song X.-Q., Chen M., Hopkins K.M., Lieberman H.B.,
RA Zhao Y.;
RT "Involvement of rhp23, a Schizosaccharomyces pombe homolog of the
RT human HHR23A and Saccharomyces cerevisiae RAD23 nucleotide excision
RT repair genes, in cell cycle control and protein ubiquitination.";
RL Nucleic Acids Res. 30:591-591(2002).
RN [2]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Reinhardt R., Pohl T.M.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Huret S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Rhode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RN Nature 415:871-880(2002).
[3]
CHARACTERIZATION.
RP MEDLINE=20119127; PubMed=10652237;
RX Lombaerts M., Goeloe J.I., den Dulk H., Brandsma J.A., Brouwer J.;
RT "Identification and characterization of the rhp23(+) DNA repair gene
RT in Schizosaccharomyces pombe.";
RL Biochem. Biophys. Res. Commun. 268:210-215(2000).
CC -!- FUNCTION: Involved in postreplication repair of UV-damaged DNA.
CC Postreplication repair functions in gap-filling of a daughter
CC strand on replication of damaged DNA.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Contains 1 ubiquitin-like domain.
CC -!- SIMILARITY: Contains 2 UBA domains.
CC
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.ebi.ac.uk/announcements
or send an email to license@ebi.ac.uk).
-----
EMBL; AF174293; AAD51975.1; -
EMBL; AL031788; CA21170.1; -
PIR; T40115; T40115.
HSSP; P54725; 1DV0.
GeneDB SPombe; SPBC2D10.12; -
InterPro; IPR004806; Rad23.
InterPro; IPR006636; STI1.
InterPro; IPR000449; UBA_Domain.
InterPro; IPR000626; Ubiquitin.
Pfam; PF00627; UBA; 2.
Pfam; PF00240; ubiquitin; 1.
SMART; SM00727; STI1; 1.
SMART; SM00165; UBA; 2.
SMART; SM00213; UBQ; 1.
TIGRFAMs; TIGR00601; rad23; 1.
PROSITE; PS50030; UBA; 2.
PROSITE; PS50053; UBIQUITIN 2; 1.
DNA damage; DNA repair; Nuclear protein; Repeat.
DOMAIN 1 77 UBIQUITIN-LIKE.
DOMAIN 135 185 UBA 1.
DOMAIN 320 360 UBA 2.
DOMAIN 119 122 POLY-ALA.

```

```
FT DOMAIN 205 208 POLY-GLN.
SQ SEQUENCE 368 AA; 4013 MW; 5CE75EB7E190EFD4 CRC64;

Query Match
Best Local Similarity 28.9%; Score 588; DB 1; Length 368;
Matches 146; Conservative 72; Mismatches 138; Indels 58; Gaps 13;

Qy 1 MKNLTKLGTNFEI-EASPDASVADVKRIETTTQGSTYRADQOQMLIYQKILKDETTL 59
Db 1 MNLTKLQOQKFVSDVSADTKISELKEIQT---QNYEVERQKLIYSGRILADKTV 57

Qy 60 ESNQVAENSLVIMLSKAKASSGASTATTAKAPATLAQAPAPVAPASVARTPTQAPVA 119
Db 58 GEYNIKEQDFIVCMVRPKTSTS-----TPKSAASPAP--NPPASVPEKKVKEAPSS 106

Qy 120 TAETAPSVQAPAPAT---VAATDDADVYQAAASNLVFGNLEQTIQIILDMGGTWE 176
Db 107 TVABESTTQTTVAASAPSNPDTTATSEAPI---DANTLAVGAQRNVAENVMENG---YE 160

Qy 177 RDTVVRALRAAYNNPERAIDLYSGIPENV---EAQPVARAPAGQQTNOQAASPAQPAV 233
Db 161 RSEVERAMRAAFNPDRAVEYLLTGIPEDILNQRSESAALAAQOQOQSEALA----- 213

Qy 234 ALPVQPSPASAGPNANPLNFPQGVPGSGSNPGVPG-AGSGALDALROLPOFQALLQLV 292
Db 214 -----PTSTG---QANLFEQAALSENENQEPSNTVGDGDPGLGFLRSIPQFQQLRQIV 263

Qy 293 QANPQIILQMLQELKQNPQILRLIOENQAEFLRLVNESPEGG---PGNILQQLAAVP 349
Db 264 QONPQMLTLLQIQGGQDPAQAQAITQNPFAFLQLAAGAESALPSSGI----- 314

Qy 350 QTLVTPTPEERAIQRLGSGNRELVEVFACNKEDELTANYLLDGHGHEFDDQ 403
Db 315 -QIQTQEESESDRLCQLGDFDRNIVIQAYLADCKNEELAANYLFEHGHSEDE 367

RESULT 4
R23A_MOUSE
ID R23A_MOUSE STANDARD; PRT; 363 AA.
AC P54726;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE UV excision repair protein RAD23 homolog A (MHR23A).
GN RAD23A OR MHR23A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chordata; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=BALB/c; TISSUE=Testis;
RX MEDLINE=86403997; Pubmed=8808275;
RA van der Spek P.J., Visser C.E., Hanaoka F., Smit B.,
RA Hagemeijer A., Bootsma D., Hoeijmakers J.H.J.;
RT "Cloning, comparative mapping, and RNA expression of the mouse
RT homologues of the Saccharomyces cerevisiae nucleotide excision repair
RT gene RAD23."
RT Genomics 31:20-27(1996).
CC -1- FUNCTION: Involved in postreplication repair of UV-damaged DNA.
CC Postreplication repair functions in gap-filling of a daughter
CC strand on replication of damaged DNA (Potential).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: Contains 1 ubiquitin-like domain.
CC -1- SIMILARITY: Contains 2 UBA domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC EMBL; X92410; CAA63145.1; -.
DR HSSP; P54725; LDV0.
DR MGD; MGI:105126; Rad23a.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006636; STIL.
DR InterPro; IPR000449; UBA domain.
DR Pfam; PF006626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00727; STIL; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBQ; 1.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS50030; UBA; 2.
DR PROSITE; PS50053; UBIQUITIN 2; 1.
KW DNA damage; DNA repair; Nuclear protein; Repeat.
FT DOMAIN 1 79 UBIQUITIN-LIKE.
FT DOMAIN 161 201 UBA 1.
FT DOMAIN 318 358 UBA 2.
SQ SEQUENCE 363 AA; 39769 MW; 67EAB968EBA5C203 CRC64;

Query Match 28.3%; Score 577; DB 1; Length 363;
Best Local Similarity 33.4%; Pred. No. 8.6e-26;
Matches 140; Conservative 77; Mismatches 124; Indels 78; Gaps 13;

Qy 3 LNVKTLKGNFIEASPDASVADVKRIETTTQGSTYRADQOQMLIYQKILKDETTLESN 62
Db 5 ITLTKLQOQTFKIRMEPDETVKVLEKEIEAKGRDAFPVAGQKLIYAGKILSDDVPIRDY 64

Qy 63 GVAENSLVIMLSKAKASSGASTATTAKAPATLAQAPAPVAP---AASVARTP----- 113
Db 65 HIDEKNFVVVMTKKA-CQGISAPPEA-SPTAVPEPSTFPFVLASGSHSPPTSRDK 122

Qy 114 TQAPVATAETAPSVQAPAPATVAATDDADVYQAAASNLVFGNLEQTIQIILDMGGG 173
Db 123 SPSEESTTTTSPESISGSPSSGSGREED-----AASLTVTGSEYETMLTEINSG-- 174

Qy 174 THERDTVVRALRAAYNNPERAIDLYSGIPENVQAPVAPAPAGQ-QTNQQAASPAQPA 232
Db 175 -YERERVAALRAASYNNPHRAVEYLLTGIPGSP-----PEHGSVQESRAEQPATEA 226

Qy 233 VALPVQPSPASAGPNANPLNFPQGVPGSGSNPGVPGAGSGALDALROLPOFQALLQLV 292
Db 227 -----AGENPLEFLRDPQFQNNRQVI 248

Qy 293 QANPQIILQMLQELKQNPQILRLIOENQAEFLRLVNESPEGGPGNI-----LGQLAA 346
Db 249 QONPALLPALQOQENPQLLQIISRHOEQFIQMLNEPP--GELADISDVEGEVGAIGE 306

Qy 347 AVPQT--LVTPTPEERAIQRLGSGNRELVEVFACNKEDELTANYLLDGHGHEFDDQ 403
Db 307 EAPQMYNIQVTPQEKAEIERLKALGFPPESLVIQAYFACEKNENLAANFULL--SQNPDDE 363

RESULT 5
R23A_HUMAN
ID R23A_HUMAN STANDARD; PRT; 363 AA.
AC P54725;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE UV excision repair protein RAD23 homolog A (HHR23A).
GN RAD23A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RX MEDLINE=94222030; Pubmed=8168482;
RA Masutani C., Sugawara K., Yanagisawa J., Sonoyama T., Ui M.,
RA Enomoto T., Takio K., Tanaka K., van der Spek P.J., Bootsma D.,
```

RA Hosijsmakers J.H.J., Hanaoka F.;
 RT "Purification and Cloning of a nucleotide excision repair complex
 RT involving the Xeroderma pigmentosum group C protein and a human
 RT homologue of yeast RAD23.";
 RL EMBO J. 13:1831-1843(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Lamerdin J.E., McCready P.M., Stilwagen S., Ramirez M., Carrano A.;
 RT "Characterization by genomic sequence analysis of a gene-rich 111 kb
 RT region of 19p13.2 containing the human DNA repair gene, RAD23A.";
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP STRUCTURE BY NMR OF 319-363.
 RX MEDLINE=99061330; PubMed=9846873;
 RA Dieckmann T., Withers-Ward E.S., Jarosinski M.A., Liu C.F.,
 RA Chen I.S.Y., Feigon J.;
 RT "Structure of a human DNA repair protein UBA domain that interacts
 RT with HIV-1 Vpr.";
 RL Nat. Struct. Biol. 5:1042-1047(1998).
 RN [5]
 RP STRUCTURE BY NMR OF 319-363.
 RX MEDLINE=20541363; PubMed=11087358;
 RA Withers-Ward E.S., Mueller T.D., Chen I.S.Y., Feigon J.;
 RT "Biochemical and structural analysis of the interaction between the
 RT UBA(2) domain of the DNA repair protein HHR23A and HIV-1 Vpr.";
 RL Biochemistry 39:14103-14112(2000).
 CC -!- FUNCTION: Involved in postreplication repair of UV-damaged DNA.
 CC Postreplication repair functions in gap-filling of a daughter
 CC strand on replication of damaged DNA (Potential).
 CC -!- SUBUNIT: Interacts with MJD.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- DOMAIN: The ubiquitin-like domain mediates interaction with MJD.
 CC -!- SIMILARITY: Contains 1 ubiquitin-like domain.
 CC -!- SIMILARITY: Contains 2 UBA domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D21235; BAA04767.1; --
 DR EMBL; AD000092; AAB51177.1; --
 DR EMBL; BC014026; AAB14026.1; --
 DR PIR; S44443; S44443.
 DR PDB; 1DVO; 11-FEB-00.
 DR PDB; 1P4I; 20-DEC-00.

DR PDB; 1IFY; 03-JUL-02.
 DR Genew; HGNC:9812; RAD23A.
 DR MIM; 600061; --
 DR GO; GO:0003697; F: single-stranded DNA binding activity; TAS.
 DR GO; GO:0006289; P: nucleotide-excision repair; TAS.
 DR InterPro; IPR004806; Rad23.
 DR InterPro; IPR006636; STI1.
 DR InterPro; IPR000449; UBA_domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 2.
 DR Pfam; PF00240; ubiquitin; 1.
 DR SMART; SM00727; STI1; 1.
 DR SMART; SM00165; UBA; 2.
 DR SMART; SM0213; UBQ; 1.
 DR TIGRFAMs; TIGR00601; rad23; 1.
 DR PROSITE; PS00303; UBA; 2.
 DR PROSITE; PS00053; UBIQUITIN_2; 1.
 KM DNA damage; DNA repair; Nuclear protein; 3D-structure; Repeat.
 FT DOMAIN 1 81 UBIQUITIN-LIKE.
 FT DOMAIN 161 201 UBA 1.
 FT DOMAIN 318 358 UBA 2.
 FT HELIX 321 324
 FT TURN 325 331
 FT HELIX 334 341
 FT TURN 342 345
 FT HELIX 348 354
 FT TURN 355 356
 SQ SEQUENCE 363 AA; 35609 MW; C4E47E9313BB47B5 CRC64;
 Query Match 28.1%; Score 573; DB 1; Length 363;
 Best Local Similarity 33.3%; Pred. No. 1.4e-25;
 Matches 139; Conservative 79; Mismatches 124; Indels 76; Gaps 13;
 QY 3 LNVKTLKGTNFEIEASPDASVADVKLIETQGSYTRADQOMLIYQKILKDETTLESN 62
 Db 5 ITLTKLQQQTFFIRMEPDETQVKVLEKIEAKGRDAPFVAGQKLIYAGKILSDDDVPIRDY 64
 QY 63 GVAENSLVIMLSKAKASSGASTATTAKAPATLAQPA--PVAPASVARTTQAPVA 119
 Db 65 RIDEKFFVVMVTKA--GQTSAPPEASPTAEPSSSTFPAPTSGHSP--PAA 118
 QY 120 TASTAP-----PSVQQAAPAAATVAATDDADVYSQAASNLVFGNNLEQTTIQIILDMGGGT 174
 Db 119 REDKSPSEASAPTPSPESV--SGSVPSGSGSREDAASTLVTGSEYETMLTEIMSG-- 174
 QY 175 WERTVTRALRAAYNNPERAIDLYLSGIPENVAQPARAPAQ--QTNQQAASPAQPAV 233
 Db 175 YERERVVAALRASYNPHRAVEYLLTGIPGSP-----PEHGSVQESQVSEQPATEA-- 226
 QY 234 ALPVQSPASAGPNANPLNLPQGVPSGSGNPGVPGAGSGALDALRQLPQFQALLQVQ 293
 Db 227 -----AGENPLEFURDQFQFQNMQRVIQ 249
 QY 294 ANPQILQPMQLQELGKONPQILRLIOENQABFLRLVNESPEGGPGGNI-----LGQLAAA 347
 Db 250 QNPALLPALLQQLQGENPOLLQIISRHQEQIQLNLEPP--GELADISDVEGEVGAIGE 307
 QY 348 VPQT--LTVTPPERAIQRLGMPNRELVLVFFACNKKDELTANYLLDHGHEFDQ 403
 Db 308 APQMWYIVTPPQEKAIERLKALGFPELSLVIAQYAFACEKNENLAANFL--SQNFDD 363
 RESULT 6
 RA23 YEAST STANDARD; PRT; 398 AA.
 ID RA23 YEAST
 AC P32628;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE UV excision repair protein RAD23.
 GN RAD23 OR YEL037C OR SYGP-ORF29.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

CC -!- SIMILARITY: Contains 1 ubiquitin-like domain.
CC -!- SIMILARITY: Contains 1 UBA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L40587; AAB07267.1; -;
DR EMBL: Z49704; CAA89774.1; -;
DR PIR: S54583; S54583.
DR HSSP: Q15843; 1NDD.
DR SGD: S0004889; DSK2.
DR InterPro: IPR006636; ST11.
DR InterPro: IPR000449; UBA domain.
DR InterPro: IPR000626; Ubiquitin.
DR Pfam: PF00627; UBA; 1.
DR Pfam: PF00240; ubiquitin; 1.
DR PRINTS: PR00348; UBIQUITIN.
DR SMART: SM00727; ST11; 2.
DR SMART: SM00165; UBA; 1.
DR SMART: SM00213; UBQ; 1.
DR PROSITE: PS50030; UBA; 1.
DR PROSITE: PS00299; UBIQUITIN_1; 1.
DR PROSITE: PS50053; UBIQUITIN_2; 1.
KW Nuclear protein.
FT DOMAIN 1 76 UBIQUITIN-LIKE.
FT DOMAIN 327 371 UBA.
FT CONFLICT 109 109 R -> A (IN REF. 2).
FT CONFLICT 296 296 R -> A (IN REF. 2).
SQ SEQUENCE 373 AA; 39516 MW; 25EDF82B9DB67DF6 CRC64;

Query Match 8.2%; Score 167; DB 1; Length 373;
Best Local Similarity 23.2%; Pred. No. 0.0085;
Matches 94; Conservative 47; Mismatches 168; Indels 96; Gaps 19;

QY 1 MSLNVTKLKGTN-FEIEASPDASVADVKRIETTGOSTYRADQOQMLYQGIKLDKDTTL 59
DB 1 MSLNHIKSGQDKWEVNVAPSTVLQFKAINKANG---IPVANORLIYSGKILKDDQTV 57

QY 60 EENGVAENSLVIMLSKAR---ASSGSASTATKAPATLAQAPAPVAPAAASVARTPTQA 116
DB 58 ESYHIQDGHSHLVKSPKPPQTASAAGANNATGCA-----AAGTGATPNWS 104

QY 117 PVATAETAPSPVQPOAAPATVAATDADVYSQAASNLVFGNNLEQTIOQLIDMGGGTWE 176
DB 105 S-QQSRGFNPLADLTSARYAGVLYNMPADMFPGDGGAL---NNDSSNNODELLRM---ME 156

QY 177 RUTVVRALRAAYNNPERADLYSGIPENVEAQPVARAPAAQQTNOQAAP-----AQA 232
DB 157 NPIFOSQNMELSNPQ-MLDFMIQSNPQLQVGPQARQLSPMPFRQMLTNPDMIRQSNQ 215

QY 233 VALPVQPSPASAGNPANPLNLPFGVPSPG-----GSPNGVVPVPGAGSGA 275
DB 216 FARWMDPN-AGHSAGGASAPF--APGDDAPEEGSNTNTSSNTGNAGTNAGTNAGA 272

QY 276 LDALRQLPQFALLQLVQANP--QILQPLMBELGKQNPQILRLIQENQAEFLRLVNESPE 333
DB 273 NTA-----ANPFASLLNPAL-----NP-----FANAGNRAST 299

QY 334 GPGGNN--ILGQLAAVPTQL-----TVTPEER--EAIQLEGMGF 370
DB 300 GMPADFDPALLASMFQPPVQASQAEDTRPPEERYEHQLRLQNDMGF 344

RESULT 8
YAUG_SCHPO STANDARD; PRT; 354 AA.
AC Q10169;
DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein C26A3.16 in chromosome I.
GN SPAC26A3.16.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidaigo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goifeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RA "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880 (2002).
CC -!- SIMILARITY: TO YEAST DSK2
CC -!- SIMILARITY: Contains 1 ubiquitin-like domain.
CC -!- SIMILARITY: Contains 1 UBA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z69240; CAA93239.1; -;
DR PIR: T38404; T38404.
DR GeneDB SPombe; SPAC26A3.16; -;
DR InterPro: IPR006636; ST11.
DR InterPro: IPR000449; UBA domain.
DR InterPro: IPR000626; Ubiquitin.
DR Pfam: PF00627; UBA; 1.
DR Pfam: PF00240; ubiquitin; 1.
DR SMART: SM00727; ST11; 2.
DR SMART: SM00165; UBA; 1.
DR SMART: SM00213; UBQ; 1.
DR PROSITE: PS50030; UBA; 1.
DR PROSITE: PS50053; UBIQUITIN_2; 1.
KW Hypothetical protein.
FT DOMAIN 1 78 UBIQUITIN-LIKE.
FT DOMAIN 309 353 UBA.
SQ SEQUENCE 354 AA; 36819 MW; 1A99B2D97E73A831 CRC64;

Query Match 7.6%; Score 154.5; DB 1; Length 354;
Best Local Similarity 21.5%; Pred. No. 0.04;
Matches 81; Conservative 58; Mismatches 178; Indels 59; Gaps 12;
DT 01-OCT-1996 (Rel. 34, Created)


```

QY 113 PTQAPVATETAAPPVQQAAPAAATVAATDDADVVSQAASNLVFGNNLEQTIQIILDMGG 172
Db 268 PPPPPAPATLARPPGEPGPTAAPAVPPPA-----AAQN-----G 304
QY 173 GTWERDVTVRALRAAYNNPERAIDLVSIGIPENVEAQP-----VARAPAAGOOTNQ----- 223
Db 305 GS-----AGAAPAPAPA-----AGGPGAGVSGQPGPGAGAAAAPAPGVKAESPVRKV 349
QY 224 QAASPAQPAVALPVQSPASA---GFNANPLMLFPQGV---PSGGSNPGVPFGAGSGALDA 278
Db 350 QAAPPAQTLAASGPASTAASWVIGPTWQGLPSPAAPVPPAPGPTGLPKGAAGAVTQS 409
QY 279 LRQLPQ-----FOALLQLVQANPQILOPMLQBLGKQNPQI---LRLIQENQAELFURLVNE 330
Db 410 LSRTPATTATSGIRATLTPTVLAPRLPQPQNPNTIQNFQPLPGMVLVRSNGQLLMIQPO 469
QY 331 SPFGPGGNGILGOLAA---AVPOTLTVTP 356
Db 470 A-----LAQMQAQAHAQPOT-TMAP 488

RESULT 10
ODP2_AZOV1
AC P10802; STANDARD; PRT; 637 AA.
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase
DE complex (EC 2.3.1.12) (E2).
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=354;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 478;
RX MEDLINE=88271330; PubMed=3292237;
RA Hansmaaijer R., Janssen A., de Kok A., Veeger C.;
RT "The dihydrolipoyltransferase component of the pyruvate
RT dehydrogenase complex from Azotobacter vinelandii. Molecular cloning
RT and sequence analysis."
RL Eur. J. Biochem. 174:593-599 (1988).
RN [2]
RP SEQUENCE OF 1-15 AND 380-415.
RX MEDLINE=88082750; PubMed=3691494;
RA Hansmaaijer R., de Kok A., Jolles J., Veeger C.;
RT "The domain structure of the dihydrolipoyl transacetylase component
RT of the pyruvate dehydrogenase complex from Azotobacter vinelandii."
RL Eur. J. Biochem. 169:245-252 (1987).
RN [3]
RP LIPOYL DOMAIN CONFORMATION.
RX MEDLINE=89052887; PubMed=3191933;
RA Hansmaaijer R., Vervoort J., Westphal A.H., de Kok A., Veeger C.;
RT "Mobile sequences in the pyruvate dehydrogenase complex, the E2
RT component, the catalytic domain and the 2-oxoglutarate dehydrogenase
RT complex of Azotobacter vinelandii, as detected by 500 MHz 1H-NMR
RT spectroscopy."
RL FEBS Lett. 240:205-210 (1988).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 381-637.
RX MEDLINE=92196586; PubMed=1549782;
RA Mattevi A., Omolova G., Schulze E., Kalk K.H., Westphal A.H.,
RA de Kok A., Hol W.G.J.;
RT "Atomic structure of the cubic core of the pyruvate dehydrogenase
RT multienzyme complex."
RL Science 255:1544-1550 (1992).
RN [5]
RP STRUCTURE BY NMR OF 1-78.
RX MEDLINE=94222112; PubMed=8068086;
RA Berg A., de Kok A., Vervoort J.;
RT "Sequential 1H and 15N nuclear magnetic resonance assignments and
RT secondary structure of the N-terminal lipoyl domain of the

```

```

RT dihydrolipoyl transacetylase component of the pyruvate dehydrogenase
RT complex from Azotobacter vinelandii."
RL Eur. J. Biochem. 221:87-100 (1994).
RN [6]
RP STRUCTURE BY NMR OF 1-78.
RX MEDLINE=97234563; PubMed=9119000;
RA Berg A., Vervoort J., de Kok A.;
RT "Three-dimensional structure in solution of the N-terminal lipoyl
RT domain of the pyruvate dehydrogenase complex from Azotobacter
RT vinelandii."
RL Eur. J. Biochem. 244:352-360 (1997).
CC -1- FUNCTION: The pyruvate dehydrogenase complex catalyzes the overall
CC conversion of pyruvate to acetyl-CoA and CO(2). It contains
CC multiple copies of three enzymatic components: pyruvate
CC dehydrogenase (E1), dihydrolipoamide acetyltransferase (E2) and
CC lipoamide dehydrogenase (E3).
CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + dihydrolipoamide = CoA + S-
CC acetyldihydrolipoamide.
CC -1- COFACTOR: THE E2 COMPONENT CONTAINS THREE COVALENTLY-BOUND LIPOYL
CC COFACTORS.
CC -1- SUBUNIT: FORMS A 24-POLYPEPTIDE STRUCTURAL CORE WITH OCTAHEDRAL
CC SYMMETRY.
CC -1- SIMILARITY: BELONGS TO THE 2-OXOACID DEHYDROGENASE FAMILY.
CC -1- SIMILARITY: Contains 3 lipoyl-binding domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X12455; CAA30987.1; ALT_INIT.
DR PIR; S01017; XXAV.
DR PDB; IEAA; 31-OCT-93.
DR PDB; IEAB; 31-OCT-93.
DR PDB; IEAC; 31-OCT-93.
DR PDB; IEAD; 31-OCT-93.
DR PDB; IEAE; 31-OCT-93.
DR PDB; IEAF; 31-OCT-93.
DR PDB; IIVU; 12-MAR-97.
DR PDB; IDPB; 20-APR-95.
DR PDB; IDPC; 20-APR-95.
DR PDB; IDPD; 20-APR-95.
DR InterPro; IPR001078; 2Oxoacid_dh.
DR InterPro; IPR006256; Acef.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR004167; E3 binding.
DR InterPro; IPR003016; Lipoyl.
DR Pfam; PF00198; 2-oxoacid_dh; 1.
DR Pfam; PF00364; biotin_lipoyl; 3.
DR Pfam; PF02817; e3 binding; 1.
DR Pfam; PF00115; 2Oxoacid_dh; 1.
DR TIGRFAMs; TIGR01348; PDHac_trf_long; 1.
DR PROSITE; PS00189; LIPOYL; 3.
KW Glycolysis; Transferase; Acyltransferase; Repeat; Lipoyl;
KW 3D-structure.
FT INIT MET 0
FT DOMAIN 1 326 LIPOYL BINDING.
FT DOMAIN 327 380 E1/E3 BINDING.
FT DOMAIN 381 637 CATALYTIC.
FT BINDING 39 39 LIPOYL (POTENTIAL).
FT BINDING 156 156 LIPOYL (POTENTIAL).
FT BINDING 261 261 LIPOYL (POTENTIAL).
FT REPEAT 1 115 LIPOYL (POTENTIAL).
FT REPEAT 116 220
FT REPEAT 221 326
FT ACT_SITE 610 610 POTENTIAL.
FT STRAND 2 5
FT STRAND 12 18
FT TURN 22 23

```


28-FEB-2003 (Rel. 41, Last annotation update)
Large proline-rich protein BAT3 (HLA-B-associated transcript 3) (G3).
BAT3 OR G3.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
TISSUE=T-cell;
MEDLINE=90192810; PubMed=2156268;
Banerji J., Sands J., Strominger J.L., Spies T.;
"A gene pair from the human major histocompatibility complex encodes
large proline-rich proteins with multiple repeated motifs and a
single ubiquitin-like domain."
Proc. Natl. Acad. Sci. U.S.A. 87:2374-2378(1990).
-i- FUNCTION: UNKNOWN.
-i- SIMILARITY: Contains 1 ubiquitin-like domain.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
EMBL; M33519; AAA35587.1; --
EMBL; M33521; AAA35588.1; --
EMBL; M33520; AAA35588.1; JOINED.
PIR; A35098; A35098.
HSSP; P02248; IUBI.
Genew; HGNC:13919; BAT3.
MIM; 142590; --
GO; GO:0003822; F:MHC-interacting protein; TAS.
InterPro; IPR000626; Ubiquitin.
Pfam; PFO0240; ubiquitin; 1.
SMART; SM00213; UBO; 1.
PROSITE; PS00299; UBIQUITIN_1; 1.
PROSITE; PS00053; UBIQUITIN_2; 1.
Repeat.
KW DOMAIN 17 77 UBIQUITIN-LIKE.
FT DOMAIN 202 207 POLY-PRO.
FT DOMAIN 242 636 4 X 29 AA APPROXIMATE REPEATS.
FT REPEAT 242 270 1.
FT REPEAT 415 443 2.
FT REPEAT 574 602 3.
FT REPEAT 608 636 4.
FT DOMAIN 657 670 POLY-PRO.
SQ SEQUENCE 1132 AA; 119504 MW; E28C8A78C3DD18 CRC64;
Query Match 6.98; Score 139.5; DB 1; Length 1132;
Best Local Similarity 23.08; Pred. No. 1;
Matches 90; Conservative 51; Mismatches 159; Indels 91; Gaps 17;
QY 1 MKNVKTLL--KGNFTEIASPDASADVVKRIETQGSYRDAQQWLIYQGIKLDETT 58
Db 17 LEVLVKTLDSTQTRFIVGAQ-----MNVKEFKHIRASVIPSKEQLIYQGVLDQDK 71
QY 59 LESNGVAENSLVIMLSKAKASS---SGASTAT-TAKAPATLAQPAAPVAPAAASVARTPT 114
Db 72 LQEVNVGGK--VIHLVERAPPQTHLPSCGSSGTGSASATHGGSPGTRPGGASVHDRNA 129
QY 115 QAPVATETAPPSVQPOAAATVAATDDADVYQAAASNLVFGNNLEQTQIQLDMGGGT 174
Db 130 NSYVMVGTFLNPS---DGSADVHINMEQAPISEPRVLVMAQHMIRDIQTLL----- 180
QY 175 WERDTVVRALRAAYNNPERAIDLY-SGIPENVEAQVAPAPAA-----GQTNQQA 226
Db 181 -----SMETLPLQCGGPGQPHSQPPPPATVPATPEVALSQTSEPYE 225
QY 227 SPAQPAVAL-----PVG-----PSPASAGNANPLNLPFGQVPSGSGNPGVVPAG 272

226 SEAPPREPMEAEVEERAPAQNPETPGPAPAGP-----TPAPETN---APNHP 271
273 SCALDALRQLPOFOALLQVQNPQILOLMQEQ-----LGKQNPQLRLIQENQASFLRLV 328
272 SPA-----EYEVLEQLERLSELRQFFLQRYEVLQAAATTDYNNHGREEDQRLI 323
329 NESPEGGPGGNILQOLAAAVFPQ---TLTVTP 356
324 NLV---GESRLGLGNTFVALSDLRCLACTP 351
RESULT 14
IDP2_PSEAE
ID ODP2_PSEAE STANDARD; PRT; 547 AA.
AC Q59638;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase
complex (EC 2.3.1.12) (E2).
GN ACEF OR ACEB OR PA5016.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PAO;
RX MEDLINE=97315227; PubMed=9171401;
RA Rae J.L., Cutfield J.F., Lamont I.L.;
RT "Sequences and expression of pyruvate dehydrogenase genes from
Pseudomonas aeruginosa."
RL J. Bacteriol. 179:3561-3571(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
opportunistic pathogen."
RL Nature 406:959-964(2000).
CC -i- FUNCTION: The pyruvate dehydrogenase complex catalyzes the overall
conversion of pyruvate to acetyl-CoA and CO(2). It contains
multiple copies of three enzymatic components: pyruvate
dehydrogenase (E1), dihydrolipoamide acetyltransferase (E2) and
lipoamide dehydrogenase (E3).
CC -i- CATALYTIC ACTIVITY: Acetyl-CoA + dihydrolipoamide = CoA + S-
acetyldihydrolipoamide.
CC -i- COFACTOR: THE E2 COMPONENT CONTAINS TWO COVALENTLY-BOUND LIPOYL
COFACTORS (BY SIMILARITY).
CC -i- SUBUNIT: FORMS A 24-POLYPEPTIDE STRUCTURAL CORE WITH OCTAHEDRAL
SYMMETRY.
CC -i- SIMILARITY: BELONGS TO THE 2-OXOACID DEHYDROGENASE FAMILY.
CC -i- SIMILARITY: Contains 2 lipoyl-binding domains.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
EMBL; U47920; AAC45354.1; --
EMBL; A5004914; AAG08401.1; --
PIR; H83018; H83018.
HSSP; P10802; LEAF.
InterPro; IPR001078; 2Oxoacid_dh.

```

DR InterPro: IPR006256; AceP
DR InterPro: IPR000889; Biotin_lipoyl.
DR InterPro: IPR004167; E3_Binding.
DR InterPro: IPR003016; Lipoyl.
DR Pfam: PF00198; 2-oxoacid dh; 1.
DR Pfam: PF00364; biotin_lipoyl; 2.
DR Pfam: PF02817; e3 binding; 1.
DR ProDom: PD001115; 2-oxoacid dh; 1.
DR TIGRFAMs: TIGR01348; PDHac_trf_long; 1.
DR PROSITE: PS00189; LIPOYL; 2.
KW Glycolysis; Transferrase; Acyltransferase; Repeat; Lipoyl;
KW Complete proteome.
FT BINDING 41 LIPOYL (BY SIMILARITY).
FT BINDING 159 159
FT ACT_SITE 520 520 POTENTIAL.
FT CONFLICT 225 225 A -> V (IN REF. 1).
FT CONFLICT 295 301 GGAGATG -> AVPAQR (IN REF. 1).
FT CONFLICT 328 329 MO -> IE (IN REF. 1).
SQ SEQUENCE 547 AA; 56709 MW; 24E15CC9A590CB4 CRC64;

Query Match 6.8%; Score 137.5; DB 1; Length 547;
Best Local Similarity 23.2%; Pred. No. 0.59;
Matches 87; Conservative 50; Mismatches 119; Indels 119; Gaps 19;

QY 41 ADQQLIYQKILKDETTLESNGVAENSEFLVIMLSKAKASS-----GASTA 87
DB 30 ADQSLTLLE-----SDKASME---IPSPKAGVVVKSIAKVGDTLKEGDEILEVEGGRQP 82
QY 88 TTAKAPATLAQPAAPVAPASVARTPTQAPVATAETAPSPVQOPAPAPATVAATDADVY 147
DB 83 AEAKAEAPAPAEAPAEAE-----PAP-----APSEKPAAPAAASVDQIKVPD 128
QY 148 SOAASNLVFGNNLEQTIQILDMGGTWERDVTVRAL---RAAYNNPERAIDLYLSGIPE 204
DB 129 SAGKANVI-----EVMYKAGDTVEADQSLITLESKASMEIPSPA-----SGVVE 173
QY 205 NV-----EAQPVAPAPAGQQTNOQAASPAQPAVALPVQSPSPASAGPNAN 249
DB 174 SVSIRKVGDEVTGDIILKLKVEGAAPAAEE---QPAAPAPAQ-AAAPAAAEQKPAAAP--- 226
QY 250 PLNLFPQGVSPGSGNSPVGAGSGALDALRQLPQFQALLQVQANPQILQPMLOELGKQ 309
DB 227 -----APAKADTPAPV---GAPSRDGAK-----VHAGPAV-RMLAREFGVE 263
QY 310 NPQIL-----RLIQENQAEFLRLVNESPEGGPGGNILGQAAAAVPTLTITPTE----- 357
DB 264 LSEVKASGPKGRLKEDVQVFYKEQLQRAKSGAGATGG---AGIPP-----IPEVDFSKF 316
QY 358 ---EREAIQRLEGMG 369
DB 317 GEVEEVAMTRLMQVG 331

```

RESULT 15

```

EFG1_CANAL
ID EFG1_CANAL STANDARD; PRT; 552 AA.
AC P43064;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Enhanced filamentous growth protein.
GN EFG1 OR EFG.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10231;
RX MEDLINE=97299875; PubMed=9155024;
RA Stoldt V.R., Sonneborn A., Leuker C.E., Ernst J.F.;
RT "Efg1p, an essential regulator of morphogenesis of the human pathogen
RT Candida albicans, is a member of a conserved class of BHLH proteins

```

```

RT regulating morphogenetic processes in fungi.";
RL EMBO J. 16:1982-1991(1997).
CC -!- FUNCTION: PUTATIVE TRANSCRIPTION FACTOR THAT STIMULATES
CC PSEUDOHYPHAL MORPHOGENESIS.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: BELONGS TO THE EFG1/PHD1/STUA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z32687; CAA83640.1; -.
DR InterPro: IPR003163; Yeast_DNA_bind.
DR Pfam: PF02292; APSES; 1.
KW Transcription regulation; Nuclear protein; DNA-binding.
FT DOMAIN 34 37 POLY-GLN.
FT DOMAIN 46 50 POLY-GLN.
FT DOMAIN 83 93 POLY-GLN.
FT DOMAIN 315 326 POLY-ALA.
FT DOMAIN 332 338 POLY-PRO.
FT DOMAIN 438 446 POLY-GLN.
FT DOMAIN 470 473 POLY-GLN.
FT DOMAIN 483 493 POLY-GLN.
FT DOMAIN 496 499 POLY-GLN.
SQ SEQUENCE 552 AA; 59564 MW; F94FD94FC2E06EB7 CRC64;

Query Match 6.8%; Score 137.5; DB 1; Length 552;
Best Local Similarity 23.9%; Pred. No. 0.6;
Matches 80; Conservative 42; Mismatches 120; Indels 93; Gaps 18;

QY 79 ASSSGA-----STATTAKAPATLAQPAAPVAPASVAR--TPTQAPVATAETA-----P 125
DB 139 SAPSGAIPSNSTSGPSQPPPLPQQQAVPIPPHVSTMQQPTFVQDTLNASTSTVGQFQP 198
QY 126 PSVQPQAAPA-----ATVAATDDAD---VYSQAASNLVFGNNL-----E 161
DB 199 PGIRPVTTTWDEKTLCYQVDANNVSVVRADNNNINGTKLNLVAQMTGRDRDGIUKS 258
QY 162 QTIOQIILDMGG-----GTWERDVTVRALRAAYNNPERAIDVLY-----SGIPEN 205
DB 259 EKVRHVVKIGSMHLKGVW--IPFERALAMAQR--EQIVDMLYPLFVRDIKRVIGTGVTPN 314
QY 206 VEAQPVAPAPAGQQTNOQAASPAQPAVALPVQP-----SPASAGPNANPLNLPFGQVPS 260
DB 315 AAA---ATAAAAATATASAPPPPPPPVAAAATTTAATAISKSSG-NGNSIS-----ATS 365
QY 261 GGSNPGVVPGAGSGALDALRQ-----LPQ---FQALLQLVQANPQIL-----Q 300
DB 366 GGSNVSGASGAGSTTSPVNTKAATAAGIPQGNYYQTNQ--QQYPOQYGVQYNAPGNQNT 423
QY 301 PMLQELGKQNPQILRLIQENQAEFLRLVNESPEGG 335
DB 424 PASQPGSTTNDQYLQQQQQQQQQQMYGYSNYQGG 458

```

Search completed: December 17, 2003, 06:17:18
Job time : 15.5265 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 06:16:40 ; Search time 41.3907 Seconds
(without alignments)
2524.996 Million cell updates/sec

Title: US-09-805-550-2

Perfect score: 2036

Sequence: 1 MKNVTKLTGTFEIEASPD.....BELTANYLLDRHGHEFDQQQ 405

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1211.5	59.5	389	10 Q9STA6	Q9sta6 lycopersico
2	1199.5	58.9	378	10 Q9FF16	Q9ff16 arabidopsis
3	1198.5	58.9	378	10 Q94C35	Q94c35 arabidopsis
4	1188.5	58.4	419	10 Q9M887	Q9m887 arabidopsis
5	1187.5	58.3	382	10 Q03990	Q03990 daucus caro
6	1127	55.4	392	10 Q40742	Q40742 oryza sativ
7	945.5	46.4	371	10 Q8LA46	Q8la46 arabidopsis
8	944.5	46.4	365	10 Q94CE9	Q94ce9 arabidopsis
9	925.5	45.5	379	10 Q03991	Q03991 daucus caro
10	840.5	41.3	367	10 Q9MA10	Q9ma10 arabidopsis
11	632	31.0	409	4 Q8WUB0	Q8wub0 homo sapien
12	576.5	28.3	362	11 Q8CAP3	Q8cap3 mus musculu
13	543	26.7	246	10 Q9S9L8	Q9s9l8 arabidopsis
14	501	24.6	341	5 Q97I35	Q97i35 dictyosteli
15	481	23.6	414	5 Q9V3W9	Q9v3w9 drosophila
16	478	23.5	414	5 Q9XZE0	Q9xze0 drosophila

17	436.5	21.4	372	5	Q23451	Q23451 caenorhabdi
18	416.5	20.5	296	11	Q8BRA9	Q8bra9 mus musculu
19	376.5	18.5	748	6	P79370	P79370 oryctolagus
20	369	18.1	389	5	Q81J58	Q81j58 plasmodium
21	365.5	18.0	343	5	Q81MB7	Q81mb7 drosophila
22	361.5	17.8	113	10	Q9SA20	Q9sa20 arabidopsis
23	317	15.6	290	5	Q9VCD5	Q9vcd5 drosophila
24	271	13.3	65	10	Q9SCA8	Q9scas lycopersico
25	243	11.9	575	10	Q9FWF5	Q9fwf5 oryza sativ
26	231.5	11.4	551	10	Q94CS1	Q94cs1 arabidopsis
27	229.5	11.3	551	10	Q9S1I8	Q9s1i8 arabidopsis
28	223	11.0	554	11	Q8C835	Q8c835 mus musculu
29	222	10.9	554	11	Q8K141	Q8k141 mus musculu
30	222	10.9	582	11	Q9JJ99	Q9jj99 rattus norv
31	214.5	10.5	523	5	Q9NIF3	Q9nif3 dictyosteli
32	213	10.5	555	11	Q8C7T4	Q8c7t4 mus musculu
33	213	10.5	582	11	Q8R317	Q8r317 mus musculu
34	210	10.3	538	10	Q9S1I9	Q9s1i9 arabidopsis
35	210	10.3	595	6	Q9SMS9	Q9sm59 bos taurus
36	209	10.3	582	11	Q9QZM1	Q9qzm1 mus musculu
37	202.5	9.9	502	5	O18672	O18672 caenorhabdi
38	201.5	9.9	484	5	Q8MM76	Q8mm76 caenorhabdi
39	198	9.7	589	4	Q81XS9	Q81xs9 homo sapien
40	197.5	9.7	596	11	Q99NB8	Q99nb8 mus musculu
41	197	9.7	589	4	Q9UMX0	Q9umx0 homo sapien
42	196	9.6	601	4	Q9NRR5	Q9nrr5 homo sapien
43	193	9.5	142	10	Q9LFI9	Q9lfi9 arabidopsis
44	193	9.5	454	5	Q8MLZ6	Q8mlz6 caenorhabdi
45	193	9.5	589	4	Q9H0T8	Q9h0t8 homo sapien

ALIGNMENTS

RESULT 1

Q9STA6 PRELIMINARY; PRT; 389 AA.
AC Q9STA6;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE RAD23 protein.
GN RAD23.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
OX NCBI_taxid=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. West virginia 106; TISSUE=fruit;
RA Lemaire-Chamley M., Petit J., Raymond P., Chevalier C.;
RT "Analysis of gene expression during early tomato fruit development by
RT mRNA differential display.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ243875; CAB51544.1; .
DR HSSP; P54725; 1DV0.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006836; ST11.
DR InterPro; IPR000449; UBA domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00727; ST11; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBQ; 1.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS00053; UBIQUITIN 2; 1.
SQ SEQUENCE 389 AA; 41508 MW; D2BAEDFOFE70778A CRC64;

Query Match 59.5%; Score 1211.5; DB 10; Length 389;
Best Local Similarity 63.0%; Pred. NO. 1.3e-66;
Matches 255; Conservative 43; Mismatches 88; Indels 19; Gaps 6;

DE Hypothetical 40.1 kDa protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Shinn P.,
RA Tracy S.E., Banh J., Bowser L., Carninci P., Chung M.K.,
RA Goldsmith A.D., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Pham P.K., Quach H.L.,
RA Sakano H., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C.,
RA Toriumi M., Yamada K., Yu G., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY037181; AAK59766.1; -;
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006636; ST11.
DR InterPro; IPR00449; UBA domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00727; ST11; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBO; 1.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 378 AA; 40075 MW; E49B591527A3EC27 CRC64;

Query Match 58.9%; Score 1198.5; DB 10; Length 378;
Best Local Similarity 62.9%; Pred. No. 86-66; 75; Indels 33; Gaps 9;
Matches 256; Conservative 43; Mismatches 43;
QY 1 MKNVTKLKTGNTFEASPDASVADVXRIETTGQSTYRADOQMLYQKILDKETTLE 60
DB 1 MKIFVTKLGSNFELEVPAKDVSVKTAIEVKG-AYEPAKQMLIHQGVKLVKDETTLE 59
QY 61 SNGVAENSFLVIMLSKAKASSGASTAT-----TAKAPATLAQPAAPVAPAAASVARTPTQA 116
DB 60 ENNVVENSFVIMLSKTKASPSGASTASAPASATQPTVATFQVS-APTASV----- 111
QY 117 PVATAETAPPSVQPOAAAPATVAATDDADVVSQAASNLVFGNNLEQTICQILDMGGTWE 176
DB 112 PVPTSGTA-----TAAAPA--TAASVQTDVYQQAASNLVAGTTLESTVQOILDMGGGSD 164
QY 177 RDTVVRLAAAYNNPERAIDYLSGIPENVEAQPVARAPAAQOQTNOQAAPQAPVALP 236
DB 165 RDTVVRLAAAFNNPERAVEVLYSGIPAEIPPPVAQAPATGEQANFLAQPOQAAA---- 221
QY 237 VQSPASAGPNANPLNLPFGQVPSGGSGVGVGAGSGALDRLQLPQFALLQVQANP 296
DB 222 --PAAATGCPNANPLNLPQGMFAA-----GAGAGNLDFLRNSHQFQALRAMVQANP 273
QY 297 QILQPMQLBELGKQNPQILRLIQENQAEFLRLVNESPEGPGGNIQLGQAAAAPQLTVP 356
DB 274 QILQPMQLBELGKQNPQLVRLIQHQADEFLRLINEPVEGEE--NVNQELEAAMPQAVTVP 331
QY 357 EREAIORLEGMGFNRELVLVEFFACNKEDELTANYLDHGHEPDDQ 403
DB 332 EREAIERLEGMGFDRAVLEVFFACNKEELAANYLDHMHFEFDQ 378

RESULT 4
Q9M887 PRELIMINARY; PRT; 419 AA.
ID Q9M887
AC Q9M887;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE Putative RAD23 (AT3g02540/F16B3_17).
GN F16B3.17.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cnv. Columbia;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.B., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC F16B3 genomic sequence";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Gibson H.A., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yu G., Yu S., Shinozaki K.,
RA Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Shinn P., Banh J., Bowser L.,
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis ORF clones";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC021640; AAF32461.1; -;
DR EMBL; AY039562; AAK62617.1; -;
DR EMBL; AY113034; AAM47342.1; -;
DR HSSP; P54725; 1DV0.
DR InterPro; IPR002965; P rich extensn.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR000449; UBA domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR PRINTS; PR01217; PRICHEXTENS.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 419 AA; 44247 MW; CAA13BC4FFEB1E25 CRC64;

Query Match 58.4%; Score 1188.5; DB 10; Length 419;
Best Local Similarity 58.4%; Pred. No. 3.7e-65;
Matches 253; Conservative 53; Mismatches 82; Indels 45; Gaps 10;
QY 1 MKNVTKLKTGNTFEASPDASVADVXRIETTGQSTYRADOQMLYQKILDKETTLE 60
DB 1 MKIFVTKLKTGTHFEIVKPEDSVVDVKKNIESVQGVADYVPAKQMLIHQGVKLVKDETTIE 60
QY 61 SNGVAENSFLVIMLSKAKASSGASTAT--TAKA-----PATLAQPAAPVAP 105
DB 61 ENKVAENSFVIMWNKSPASAAASAGTSQAKSIPPTSPSTQPSISQTPASVAPVAP 120
QY 106 AAS----VARTPTQAPVATETAPPSPVQPAAPATVAAT-----DDADVVSQAAS 152
DB 121 APTPPPPAPTPTAPVAATETVTPPI-PEVPATISSSTPAPDSAPVSGQDVGQAAS 179
QY 153 NLVFGNNLEQTICQILDMGGTWERDVTVRLAALRAAYNNPERAIDYLSGIPENVEAQPVA 212
DB 180 NLAAGSNLESTICQILDMGGGTMDRETIVLALRAAFNNPERAVEVLYTGIPEQAEVPPVA 239

Query Match 58.4%; Score 1188.5; DB 10; Length 419;
Best Local Similarity 58.4%; Pred. No. 3.7e-65;
Matches 253; Conservative 53; Mismatches 82; Indels 45; Gaps 10;

Db	174	DRTVTIRIVRAAFNPERAVEVILYSGIPEQAEPVAPSPSG-----QAANPLDQPPAA	228
Qy	235	LPVQSPASAGPNANPLNFPQGVPSGGSNPGVVGAGSGALDALRQLPQFQALLQLVQA	294
Db	229	--AQPASAGPNANPLDLFPQGLPDWGSN-----AAGAGNLDFLRNQQFQALRAMVQS	281
Qy	295	NPOILQPMLOELGKQNPQILRLIQENQAEFLRLVNESPEGGPGGNILGQLAAAVPQTLTV	354
Db	282	NPOILQPMLOELGKQNPQILRLIQENQAEFLRLVNESPEGGPGGNILGQLAAAVPQTLTV	354
Qy	355	TPPERAIOLEGMGNFRELVLVFFVACNKKDEBLTANYLLDHGHEFD	401
Db	336	TPPERAIOLEGMGNFRELVLVFFVACNKKDEBLTANYLLDHGHEFE	382
RESULT 6			
Q40742			
ID	Q40742	PRELIMINARY;	PRT; 392 AA.
AC	Q40742;		
DT	01-NOV-1996 (TEMBLrel. 01, Created)		
DT	01-NOV-1996 (TEMBLrel. 01, Last sequence update)		
DT	01-MAR-2003 (TEMBLrel. 23, Last annotation update)		
DE	OSRAD23.		
OS	Oryza sativa (Rice).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
OC	Ehrhartoideae; Oryzaceae; Oryza.		
ON	NCBI_TaxID=4530;		
RX	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Nipponbare;		
RX	MEDLINE=37369378; PubMed=9225866;		
RA	Schultz T.F., Quatrano R.S.;		
RT	"Characterization and expression of a rice RAD23 gene.";		
RL	Plant Mol. Biol. 34:557-562(1997).		
DR	EMBL; U63530; AB65841.1; -.		
DR	HSP; P54725; IDV0.		
DR	Gramene; Q40742; -.		
DR	InterPro; IPR004806; Rad23.		
DR	InterPro; IPR006636; ST11.		
DR	InterPro; IPR000449; UBA domain.		
DR	InterPro; IPR000626; Ubiquitin.		
DR	Pfam; PF00627; UBA; 2.		
DR	Pfam; PF00240; ubiquitin; 1.		
DR	SMART; SM00727; ST11; 1.		
DR	SMART; SM00165; UBA; 2.		
DR	SMART; SM00213; UBQ; 1.		
DR	TIGRFAMs; TIGR00601; rad23; 1.		
DR	PROSITE; PS0053; UBIQUITIN_2; 1.		
SQ	SEQUENCE 392 AA; 41754 MW; BD6E08574CC7CACB CRC64;		
Query Match	55.4%;	Score 1127; DB 10; Length 392;	
Best Local Similarity	59.1%;	Pred. No. 2e-61;	
Matches	243; Conservative	48; Mismatches	90; Indels 30; Gaps 9;
Qy	1	MKLNKTKLKGTFNFEIEASPDASVADVKRIIETTTQGGSTYRADQOMLIYQGIKLDKDTTLE	60
Db	1	MKLSVTKLGSTFQIEVDSAQKADVVKRIIETTTQGGHIYPAEQOMLIHQGVKDKDTTLD	60
Qy	61	SNQVAENSLFVIMLSKAKASSSGASTATTAKAPATLAQP--AAPVAPASVARTPTQAPV	118
Db	61	ENKVLNSFLVIMLRQKGSS--SAPATSKAPSNQAPPTQTVPAPAA-----SQAPV	111
Qy	119	ATAETAP-----PSVQPPAPAAATVAATDDADVVSQAASNLVFGNNLEQTIQIILLDMGGG	173
Db	112	APATTVPTVTSAPPTTATAPAPAVSSADNYGQATSNLVAGSNLEATTIQSIEMGGG	171
Qy	174	TWERDVTVVRALRAAYNNPERAIDLYISGIPENVEAQPVARAPAAQQTN--QQAASPAQPA	232
Db	172	IWDREDIVLHALSAAFNNPERAVEVILYSGVPEQMDI-PV--PPPSIQANPTQASQATQPA	228
Qy	233	VALPVQSPASAGPNANPLNFPQGVPSGGSNPGVVGAGSGALDALRQLPQFQALLQLV	292

Db 229 A-----PSILSGFNASFELDFPQALPNASTD-----AAGLNLDAALRNNAQFRTLLSLV 278

Qy 293 QANPQILQMLQELGKQNPQILRLIQENQAFRLVLNVPESPGPGGNTLGLAAAVPQTL 352

Db 279 QANPQILQMLQELGKQNPQILRLIQENQAFRLVLNVPESPGPGGNTLGLAAAVPQTL 338

Qy 353 TVTPPERAIQRLGCMGFNRELVLVFFACNKKDELTANYLLDHGHEFDDQ 403

Db 339 AVTPPEDEAILRLPMGFDRALVLDVFFACNKKDELTANYLLDHGHEFDDQ 389

RESULT 7

Q8LA46 PRELIMINARY; PRT; 371 AA.

ID Q8LA46

AC Q8LA46

DT 01-OCT-2002 (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE DNA repair protein RAD23, putative.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurooids II; Brassicales; Brassicaceae; Arabidopses.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Haas B.J., Volkovskiy N., Town C.D., Troukhan M., Alexandrov N.,

RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;

RT "Full-length messenger RNA sequences greatly improve genome

RT annotation.";

RL Genome Biol. 0:0-0(2002).

RN [2]

RP SEQUENCE FROM N.A.

RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,

RA Feldmann K.;

RT "Full-length cDNA from Arabidopsis thaliana.";

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY088037; AAM65583.1; -

DR InterPro; IPR004806; Rad23.

DR InterPro; IPR006636; STI1.

DR InterPro; IPR000449; UBA domain.

DR InterPro; IPR000626; Ubiquitin.

DR Pfam; PF00627; UBA; 2.

DR SMART; SM00727; STI1; 1.

DR SMART; SM00213; UBQ; 1.

DR TIGRfams; TIGR00601; rad23; 1.

DR PROSITE; PS0053; UBIQUITIN 2; 1.

SQ SEQUENCE 371 AA; 39747 MW; 081493086EA976E7 CRC64;

Query Match 46.4%; Score 945.5; DB 10; Length 371;

Best Local Similarity 50.7%; Pred. No. 2.5e-50;

Matches 208; Conservative 60; Mismatches 95; Indels 47; Gaps 11;

Qy 1 MKNVTKLKGTFNFIASPDASVADVKRIIETQGSTYRADQMLIYQGLKDETTLE 60

Db 1 MKNVTKLKGTFNFIASPDASVADVKRIIETQGSTYRADQMLIYQGLKDETTLE 60

Qy 61 SNGVAENSLVIMLSKAKA-SGASSTATTAKAPATLAQPAAPVAPASVARTPTQAPVAT 120

Db 61 ENKVTTEGLVLMVLSKSK-SGASSTATTAKAPATLAQPAAPVAPASVARTPTQAPVAT 111

Qy 121 AETAPSVQV---QAAPATVAATDADVYSQAASNLVFGNNLEQTIQOILDMGGGTWER 177

Db 121 TQSSVPVAPSPIPAQSQPAA-----QTDYTGAAATLVSGSLSEQVMVQOIMEMGGGMDK 165

Qy 178 DTWVTRALRAYNNPERAIDYLSGIPENVE-AQPVARAPAAQQTNQQAASPAQPAVALP 236

Db 166 EIVTRALRAYNNPERAIDYLSGIPENVE-AQPVARAPAAQQTNQQAASPAQPAVALP 212

Qy 237 VQPSFASAGPNANPLNLPQGVPSGSGNPGVVPVAGS---GALDARQLPQFALLQLVQ 293

Db 213 V--APASGPGNSPLDLFPQ-----ETVAAGSGDLGLTFLEFLRNNDQFQQLRTWVH 261

Qy 294 ANPQILQMLQELGKQNPQILRLIQENQAFRLVLNVPESPGPGGNTLGLAAAVPQTL 352

Db 262 SNPQILQMLQELGKQNPQILRLIQENQAFRLVLNVPESPGPGGNTLGLAAAVPQTL 321

Qy 353 TVTPPERAIQRLGCMGFNRELVLVFFACNKKDELTANYLLDHGHEFDDQ 402

Db 322 NVTPAEQAIQRLGCMGFNRELVLVFFACNKKDELTANYLLDHGHEFDDQ 371

RESULT 8

Q94CE9 PRELIMINARY; PRT; 365 AA.

ID Q94CE9

AC Q94CE9

DT 01-DEC-2001 (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE Putative RAD23 protein.

GN F20B17.8.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurooids II; Brassicales; Brassicaceae; Arabidopses.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,

RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,

RA Bowser L., Carninci P., Chen H., Cheuk R., Hayaishizaki Y., Ishida J.,

RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E.,

RA Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,

RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,

RA Shinzaki K., Davis R.W., Ecker J.R., Theologis A.;

RT "Full length cDNA of gene F20B17.8 (GI:7715605).";

RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,

RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,

RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,

RA Hayaishizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,

RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,

RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,

RA Shinn P., Southwick A., Shinzaki K., Davis R.W., Ecker J.R.,

RA Theologis A.;

RT "Full length cDNA of gene F20B17.8 (GI:7715605).";

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY034912; AAK59419.1; -

DR EMBL; AY063103; AAL34277.1; -

DR InterPro; IPR004806; Rad23.

DR InterPro; IPR006636; STI1.

DR InterPro; IPR000449; UBA domain.

DR InterPro; IPR000626; Ubiquitin.

DR Pfam; PF00627; UBA; 2.

DR Pfam; PF00240; ubiquitin; 1.

DR SMART; SM00727; STI1; 1.

DR SMART; SM00165; UBA; 2.

DR SMART; SM00213; UBQ; 1.

DR TIGRfams; TIGR00601; rad23; 1.

DR PROSITE; PS0053; UBIQUITIN 2; 1.

SQ SEQUENCE 365 AA; 39157 MW; 4FED9EC59B467745 CRC64;

Query Match 46.4%; Score 944.5; DB 10; Length 365;

Best Local Similarity 50.5%; Pred. No. 2.8e-50;

Matches 206; Conservative 56; Mismatches 97; Indels 49; Gaps 9;

Qy 1 MKNVTKLKGTFNFIASPDASVADVKRIIETQGSTYRADQMLIYQGLKDETTLE 60

Db 1 MKNVTKLKGTFNFIASPDASVADVKRIIETQGSTYRADQMLIYQGLKDETTLE 60

Qy 61 SNGVAENSLVIMLSKAKA-SGASSTATTAKAPATLAQPAAPVAPASVARTPTQAPVAT 119

Db 61 ENKVTTEGLVLMVLSKSKSGGAGSQVQVPSATTSSTKPAAP-----STTQSSPV 111


```

QY 120 TAETAPPSVQPAAPAAATVADDDADVYSQAASNLVFGNNLEQTIQOILDMGGGTWERDT 179
DB 112 PASPIPAEQEPAAR-----QTDYTGAAASTLVSGSSLEQVMQVQIMEGGSGWDKET 161
QY 180 VTRALRAAYNNPERAIDYLSGIPENVE--AQPVARAPAAAGQQTNOQAASPAQPAVALPVQ 238
DB 162 VTRALRAAYNNPERAVDLYSGIPQTAEEVAVPVPEAQIAG-----SGAAPV-- 207
QY 239 PSASAGPNANPLNLPQGVPSGGSNPGVVPAGS---GALDALRQLPQFQALLQVQAN 295
DB 208 -APASGPNPSPLDLFPFQ-----ETVAAGSGDLGTLFLRNNDFQQLRTVHSHN 257
QY 296 PQILQPMQLQKQNPQILRIQENQAEFLRLVNESPEGGPG--GNILGQLAAAVPQTLTV 354
DB 258 PQILQPMQLQKQNPQILRIQENQAEFLRLVNEPYEGSGDGEDMFDQEQEMPHAINV 317
QY 355 TPEREAIQRLGEGFNRLEVLVFFACNCKDEBELTANYLLDHGHEFDD 402
DB 318 TPAEQEAIQRLGEGFNRLEVLVFFACNCKDEBELTANYLLDHGHEFDD 365

RESULT 9
O03991
ID O03991 PRELIMINARY; PRT; 379 AA.
AC O03991;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE RAD23 protein, isoform II.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Apiales; Apiaceae; Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W001C;
RX MEDLINE=98345997; PubMed=9681019;
RA Sturm A., Leinhardt S.;
RT "Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in yeast.";
RL Plant J. 13:815-821(1998).
DR EMBL; Y12014; CAAT2742.1; -.
DR HSP; P54725; IDV0.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006636; STII.
DR InterPro; IPR000449; UBA domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00727; STII; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBQ; 1.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS50053; UBIQUITIN 2; 1.
SQ SEQUENCE 379 AA; 40530 MW; B26697B439CC5929 CRC64;

Query Match 45.5%; Score 925.5; DB 10; Length 379;
Best Local Similarity 50.2%; Pred. No. 4.3e-49;
Matches 203; Conservative 57; Mismatches 117; Indels 27; Gaps 8;

QY 1 MKNLVKTLKGNFIEASPDASVADVRIETTGQSTYRADQOQLYQKILKDETTLE 60
DB 1 MKLTVKTLKGSFFIRAPQNDNTVMAIKNIEDLQKDNYPGCGQLLHNGKVLKDESTLA 60
QY 61 SNGVAENFLVIMLSKAKA--SSSGASTATTAKAPATLAQAPVAPAAASVARTPTQAPVA 119
DB 61 ESKISEDGLVVMVGLKSKTNSSTGTTPAAQSSASAPTPAPAPAPAPAPASAVTPNT 120
QY 120 TAETAPPSVQPAAPAAATVADDDADVYSQAASNLVFGNNLEQTIQOILDMGGGTWERDT 179
DB 121 TVPEAP--LSPAPAP-----SDTYGEAASNVVAGSNLEQTIQIHMDGGGMDTNN 169

```

```

QY 180 VVRALRAAYNNPERAIDYLSGIPENVEAQ--PVARAPAAAGQQTNOQAASPAQPAVALPVQ 238
DB 170 VSRALRAAYNNPERAVDLYSGIPMAEAAVPVSH--FGDQINAGNNAISDNGVA---- 223
QY 239 PSASAGPNANPLNLPQGVPSGGSNPGVVPAGSGALDALRQLPQFQALLQVQANPOI 298
DB 224 -GAAPGAPNSLPLNFPQETLSG-----VTGAGLGSLEFLRNNPQFQTLRSMVORNPQI 276
QY 299 LQPMQLQKQNPQILRIQENQAEFLRLVNESPEGGPGGNILGQLAAAVPQTLTVTPPE 358
DB 277 LQPMQLQKQNPQILRIQEHHEFLQLINEPVEASE--GDMFDQEQDVPQETITVTAAD 335
QY 359 REAIQRLGEGFNRLEVLVFFACNCKDEBELTANYLLDHGHEFDD 402
DB 336 QEAIRLEAMGFDRLVIEAFLACDRNEELAVNYLLENAGDFED 379

RESULT 10
Q9MA10
ID Q9MA10 PRELIMINARY; PRT; 367 AA.
AC Q9MA10;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE F20B17.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Khan S., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Kim C.,
RA Shinn P., Altafi H., Bei Q., Chin C., Chou J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howng B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharshy N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F20B17 from chromosome 1."
RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howng B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharshy N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC010793; AAF68123.1; -.
DR HSP; P54725; IDV0.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR000449; UBA domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; TIGR00601; rad23; 1.
DR PROSITE; PS50053; UBIQUITIN 2; 1.
SQ SEQUENCE 367 AA; 39790 MW; C6D01134B5C69CE6 CRC64;

```

```

Query Match      41.3%; Score 840.5; DB 10; Length 367;
Best Local Similarity 47.4%; Pred. No. 6.7e-44;
Matches 192; Conservative 53; Mismatches 119; Indels 41; Gaps 8;

QY 1 MKNVKTLLKGNTEFEASPDASVADVVKRIETTGQSTYRADQOQMLLYOGKILKDETTLE 60
DB 1 MKNVKTLLKGNTEFEASPDASVADVVKRIETTGQSTYRADQOQMLLYOGKILKDETTLE 60
QY 61 SNGVAENSFLVIMLSKAKASSGASTATTAKAPATLAQAPAPVAPAAASV 109
DB 61 SNGVAENSFLVIMLSKAKASSGASTATTAKAPATLAQAPAPVAPAAASV 109
QY 110 ARTPTQ-----APVATAETAPPSVQOQAPAAAT-VAATDD-----ADVYSQ 149
DB 120 TSTPASITPASATASSEPAPASAAKQEPKPAETPVATSPATSTDSGSSSNLFPED 179
QY 150 AASNLVFGNNLEOTIQOILDMGGGTWERTVVRALRAAYNNPERAIDLYSGIPENVEAQ 209
DB 180 ATVALVTGQSYENMVTEIMSG---YEREQVIAALRASFNPNPDRAVEYLLWIGIPGDRSQ 236
QY 210 PVARAPAAQOQTNQOAS---PAQPAVALPVQPSPASAGNPANLENLPPQGVPGGSGNP 266
DB 237 AVDDPP-----QAASTGVPOSSAVAAAAATTTATT-----TTTSSGGHP- 275
QY 267 VPGAGSGALDRLQFQALLQLOVANPOILQPMQLGKONPOILRLIQENQABFLR 326
DB 276 -----LEFLRNQFQFQOMRQIIQONPSLLPALLQOIGRENPOQLQISQOEHTIQ 326
QY 327 LVNE-----SPEGPGGNILGQLAAAVP---QTLVTTPPEERAIQRLGGMGFNRLVLE 378
DB 327 MLNEPVQEGAGQGGGGGGGGIAEAGSGHNVYIQTVPQEKAIERLKAIGFPEGLVQA 386
QY 379 FPACNKBELTANYLLDHGHEFD 401
DB 387 YFACENKLENLAANFLQONFED 409

RESULT 12
Q8CAP3 PRELIMINARY; PRT; 362 AA.
ID Q8CAP3
AC Q8CAP3;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE RAD23a homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RC MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK038300; BAC29962.1; -.
SQ SEQUENCE 362 AA; 39635 MW; 804B2608ECA241A9 CRC64;

Query Match      28.3%; Score 576.5; DB 11; Length 362;
Best Local Similarity 33.0%; Pred. No. 9.8e-28;
Matches 138; Conservative 80; Mismatches 123; Indels 77; Gaps 13;

QY 3 LNVKTLKGNTEFEASPDASVADVVKRIETTGQSTYRADQOQMLLYOGKILKDETTLESN 62
DB 5 ITLTLOOQTFKIRMEPDETIVKLEKIEAKGADFPVAGQKLIYAGKILSDDDVPKBY 64
QY 63 GVAENSFLVIMLSKAKASSGASTATTAKAPATLAQAPAPVAP--AASVARTP----- 113
DB 65 HIDEKNFVVMVTKAKAQ--GIPAPPEASTAVPEPSTPPFVLAGSMSPPTSRDK 122
QY 114 TQAPVATAETAPPSVQOQAPAAATVAATDADVYSQAASNLVFGNNLEQTTQOILDMGGG 173
DB 123 SPSESTTTTSPESIGSVSPSSGSGREED-----AASLTVTGSEVETMLTEIMSG-- 174
QY 174 TWEDTIVRALRAAYNNPERAIDLYSGIPENVEAQPVARAPAAQOQTNQOASPAQAV 233
DB 175 -YERERVAALRASNNPHRAVEYLLTGIGSPS-----PEHG--SVQSOAPEQAT 224

Query Match      41.3%; Score 840.5; DB 10; Length 367;
Best Local Similarity 47.4%; Pred. No. 6.7e-44;
Matches 192; Conservative 53; Mismatches 119; Indels 41; Gaps 8;

QY 1 MKNVKTLLKGNTEFEASPDASVADVVKRIETTGQSTYRADQOQMLLYOGKILKDETTLE 60
DB 1 MKNVKTLLKGNTEFEASPDASVADVVKRIETTGQSTYRADQOQMLLYOGKILKDETTLE 60
QY 61 SNGVAENSFLVIMLSKAKASSGASTATTAKAPATLAQAPAPVAPAAASV 109
DB 61 SNGVAENSFLVIMLSKAKASSGASTATTAKAPATLAQAPAPVAPAAASV 109
QY 110 ARTPTQ-----APVATAETAPPSVQOQAPAAAT-VAATDD-----ADVYSQ 149
DB 120 TSTPASITPASATASSEPAPASAAKQEPKPAETPVATSPATSTDSGSSSNLFPED 179
QY 150 AASNLVFGNNLEOTIQOILDMGGGTWERTVVRALRAAYNNPERAIDLYSGIPENVEAQ 209
DB 180 ATVALVTGQSYENMVTEIMSG---YEREQVIAALRASFNPNPDRAVEYLLWIGIPGDRSQ 236
QY 210 PVARAPAAQOQTNQOAS---PAQPAVALPVQPSPASAGNPANLENLPPQGVPGGSGNP 266
DB 237 AVDDPP-----QAASTGVPOSSAVAAAAATTTATT-----TTTSSGGHP- 275
QY 267 VPGAGSGALDRLQFQALLQLOVANPOILQPMQLGKONPOILRLIQENQABFLR 326
DB 276 -----LEFLRNQFQFQOMRQIIQONPSLLPALLQOIGRENPOQLQISQOEHTIQ 326
QY 327 LVNE-----SPEGPGGNILGQLAAAVP---QTLVTTPPEERAIQRLGGMGFNRLVLE 378
DB 327 MLNEPVQEGAGQGGGGGGGGIAEAGSGHNVYIQTVPQEKAIERLKAIGFPEGLVQA 386
QY 379 FPACNKBELTANYLLDHGHEFD 401
DB 387 YFACENKLENLAANFLQONFED 409

RESULT 11
Q8WUB0 PRELIMINARY; PRT; 409 AA.
ID Q8WUB0
AC Q8WUB0;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE RAD23 homolog B (S. cerevisiae).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020973; AAH20973.1; -.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006836; STII.
DR InterPro; IPR000449; UBA domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00727; STII; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBQ; 1.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS0053; UBIQUITIN 2; 1.
SQ SEQUENCE 409 AA; 43199 MW; 475FBD499DACAC69 CRC64;

Query Match      31.0%; Score 632; DB 4; Length 409;
Best Local Similarity 34.5%; Pred. No. 4.5e-31;
Matches 153; Conservative 81; Mismatches 133; Indels 76; Gaps 12;

QY 1 MKNVKTLLKGNTEFEASPDASVADVVKRIETTGQSTYRADQOQMLLYOGKILKDETTLE 60
DB 1 MKNVKTLLKGNTEFEASPDASVADVVKRIETTGQSTYRADQOQMLLYOGKILKDETTLE 60

```

```
QY 234 ALPVQSPASAGNPANPLNLPFGQVPSGSGNSPGVPGAGSGALDRLQLPQFQALLQLVQ 293
Db 225 -----EAGENP-----LEFLRDQPFQFQMFQVQ 248
QY 294 ANPQILQPMQLQELGKQNPQLRLIQENQAEFLRLVNESPEGGPGNI-----LGLAAA 347
Db 249 QNPALLPQLOQLGOENPQLQQLSRHQEQFIQMLNEPP--GELADISDVEGEVGAIGSE 306
QY 348 VPQT--LTVTPEREAIQRLGEGFNRRELVLVFFACNKGDELTANYLLDHGHEFDQ 403
Db 307 APQMNVIQVTPQSEAIERLKLGFPELSVQIAYFACEKNENLAANFL--SQNFDE 362

RESULT 13
Q9S9L8 PRELIMINARY; PRT; 246 AA.
AC Q9S9L8;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE T24D18.27 protein.
GN T24D18.27.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Liu S.X., Yu G., Sakano H., Jhaveri A., Lee J.M., Lenz C., Pham P.,
RA Toriumi M., Chin C., Chio J., Choi E., Chung M., Gonzalez A.,
RA Howing B., Koo T., Li J., Liu A., Vaysberg M., Altafi H., Brooks S.,
RA Buehler E., Chao Q., Conn L., Conway A., Hansen N., Johnson-Hopson C.,
RA Khan S., Kim C., Lam B., Nguyen M., Palm C., Shinn P., Tambunga G.,
RA Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;
RA "The sequence of BAC T24D18 from Arabidopsis thaliana chromosome 1.";
RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RL 2;
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Theologis A.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR HSSP; P02248; IUBI.
DR InterPro; IPR000449; UBA domain.
DR Pfam; PF00627; UBA; 1.
DR SMART; SM00165; UBA; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS50053; UBIQUITIN 2; 1.
SQ SEQUENCE 246 AA; 25966 MW; 0867D1E7D72B6FPE CRC64;

Query Match 26.7%; Score 543; DB 10; Length 246;
Best Local Similarity 41.8%; Pred. No. 6.9e-26;
Matches 117; Conservative 41; Mismatches 74; Indels 48; Gaps 4;

QY 1 MKNLVKTLKGTNFEIEASPDASVADVVKRIETTGQSTYRADQQLIYQKILKDETTLE 60
Db 1 MKNLVKTLKGSHEIRVLPDTTMAVKNIEDSQKDNYPGQQLIHNGKVLKDETTLV 60
QY 61 SNGVAENSLVIMLSKAKASSGASTATTAKAPATLAQPAAPVAPASVARTTQAPVAT 120
Db 61 ENKVTGEGFLVIMLSKSKTASSAGSPSTQLAARSTTQSIAPASNSTPVEQETA----- 115
QY 121 AETAPPSVQPAAPATAATDADVYSQAASNLVFGNNLEQTIQILDGCGGTWERDVT 180
Db 116 -----QSDTYGQAASLTLSGSSSTEQVMVQIMENGSGSWDKETV 153
QY 181 VRLRAAYNNPERAIDYLSGIPENVEAQVAPAPAGQTNQQAASPAQPAVALPVQPS 240
Db 154 TRALRAAYNNPERAVDYLSGIPETV---TIPATNLSGVSGSRELTAP----- 198
```

```
QY 241 PASAGNPANPLNLPFGQVPSGSGNSPGVPGAGSGALDRL 280
Db 199 PPSGGPNSPLDFPQEAIVSDAA-----GGDLGTLEFLR 232

RESULT 14
Q97135 PRELIMINARY; PRT; 341 AA.
AC Q97135;
DT 01-MAY-1999 (T-EMBLrel. 10, Created)
DT 01-MAY-1999 (T-EMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE RepC-binding protein A.
GN RCBA.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=slb153;
RA Li G., Alexander H., Alexander S.;
RT "rcba, the Dictyostelium discoideum homolog of yeast repair gene RAD23.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF103870; AAD17913.1; -.
DR HSSP; P54725; IDVO.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006636; STI1.
DR InterPro; IPR000449; UBA domain.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00727; STI1; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBQ; 1.
DR TIGRfam; TIGR00601; rad23; 1.
DR PROSITE; PS50053; UBIQUITIN 2; 1.
SQ SEQUENCE 341 AA; 37528 MW; E376B909E6DEB57E CRC64;

Query Match 24.6%; Score 501; DB 5; Length 341;
Best Local Similarity 31.9%; Pred. No. 3.9e-23;
Matches 136; Conservative 59; Mismatches 123; Indels 108; Gaps 15;

QY 1 MKNLVKTLKGTNFEIEASPDASVADVVKRIETTGQSTYRADQQLIYQKILKDETTLE 60
Db 1 MKVTIKNKKEIYVFEVNGDLTVAEIKNLISEKHQT---PSWQTLIYSGKILEDKRTLE 57
QY 61 SNGVAENSLVIMLSKAKASSGASTATTAKAPATLAQPAAPVAPASV----- 109
Db 58 SYNITDSGFIKXMIKKPR-----EAPAT--TPAPSTTPLNLTSTNHHHYCRT 103
QY 110 -----ARTPTQAPVATAETAPPSVQPAAPATAATDADVYSQAASNLVFGN 158
Db 104 NNHKLNTNNTTTTPTTPTTNT--PATFNPPTTSTPGSTSTTS--POOSSDFATGT 160
QY 159 NLEQITQIQLDMGGGTWERDVTVRLRAAYNNPERAIDYLSGIPENVEAQVAPAPAA 218
Db 161 ELEATIKNTDNG---FARDQVLRALRLTFNNAERAIEVLVSG-----NIPAA 206
QY 219 QQTNOQAASPAQPAVALPVQSPASAGNPANPLNLPFGQVPSGSGNSPGVPGAGS 278
Db 207 DPEDBEE-----MEGGG-----GSGDNPFPA 227
QY 279 LRQLPQFQALLQLVQNPQLQPMQLQELGKQNPQLRLIQENQAEFLRLVNESPEGGPG 338
Db 228 LRNHFNHLLRAISKNPSIIFGILOQLAQTNPALVRQIQENPNEFIRLF--QGDNPG 285
QY 339 NILGOLAAVPTLTTPEREAIQRLGEGFNRRELVLVFFACNKGDELTANYLLDHG 397
Db 286 N-PGQF-----TLQVTQESEAIRQLQALTGMDKSTVIEAFACDKNELTASYL 338
QY 398 HEFDQ 403
```

```

Db ||:
339 ---DDE 341

RESULT 15
Q9V349 PRELIMINARY; PRT; 414 AA.

AC Q9V3W9;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE RAD23 protein.
GN RAD23 ORF C61836.

OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_taxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.P.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Buesam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Fowler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Hernandez J.R., Houck J.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
[2]
RN SEQUENCE FROM N.A.
RP Brodey M.H., Rubin G.M., Tsang G.;
RT "Full Length Drosophila melanogaster cDNA sequence.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003844; AAF59352.1; -
DR EMBL; AF132147; AAD33594.1; -
DR HSSP; P54725; 1DV0.
DR FlyBase; FBgn026777; Rad23.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006636; Rad23.
DR InterPro; IPR000449; UBA_domain.
DR InterPro; IPR000626; Ubiquitin.

```

```

DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00727; STI1; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBQ; 1.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS00553; UBIQUITIN_2; 1.
SQ SEQUENCE 414 AA; 45780 MW; 4C2E494CA116F7AB CRC64;

Query Match 23.6%; Score 481; DB 5; Length 414;
Best Local Similarity 29.3%; Pred. No. 8.2e-22;
Matches 134; Conservative 72; Mismatches 153; Indels 98; Gaps 10;

Qy 1 MGLNVKTLKGTNFEIEASPDASVADVADVKRIITQQGQSTYRADQOQMLYQGIKILDETTLE 60
Db 1 MIITIKNLQOQTFTIEPAPEKTVLELKKKIFEEGRPE-YVAEKQLIYAGVILTDDRTVG 59
Qy 61 SNGVAENSFLVIMLSKAKASS-----GASTATTAKAPATLAQPAAPAPAPASVARTPT 114
Db 60 SYNVDKKFTVVMVLRDSSSSNRNQLSVKESNKLTTDDSKSQSMPCPEANHTNPSSTNT 119
Qy 115 QAPVATAETAPPSPVQQAAPAAVTAATDDADVVSQAASNLVFGNNLEQTIIQIILDMGGT 174
Db 120 EDSVLSRETRPLSSDELIGELA-----QASLOSRAESNLMDGTYNQTIVLSMVEMG--- 170
Qy 175 WERTVTVRALRAAYNNPERRAIDYLYSGIPENVEAQPVARAPAAAGQOQNOQAASPAQPAVA 234
Db 171 YPREQVERAASVNNPRAVEVLINGIP-----AEEGTFYRNLESTNPSL 217
Qy 235 LPVQSPASAGPNANPLNLPFGVPGSGSNPGVVGAGSGALDALROLPOFQALLQLVQA 294
Db 218 IPGSPQPPASA-----TSAERSTESN-----SDPFBLRSQPFQWRSLIYQ 259
Qy 295 NPQILQMLQELGKQNPQILRLIOENQAELRLVNE--SPEGPGGNNILQQLAAAVPQT 352
Db 260 NPHLLHAVLQOIGQTNPALLQILISENQDAFLNMLNQPIDRESSESGATVPVPSNARIFSTL 319
Qy 353 -----TVTPEEREAIQRL 365
Db 320 DNVDLFSPDLEVATSAQSAQTSAAHQSGSAADNEDLEQPLGVSTIRLNQDKDAIERL 379
Qy 366 EGMGFNRELVLVFFACNKKDELTANYLLDHGHEFDD 402
Db 380 KALGFPEALVLQAYFACEKNEQAANFL--SSSFDD 414

Search completed: December 17, 2003, 06:20:06
Job time : 45.3907 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 06:16:39 ; Search time 38.771 Seconds
(without alignments)
1658.049 Million cell updates/sec

Title: US-09-805-550-2

Perfect score: 2036

Sequence: 1 MKNVKTLLKGNFIEASPD.....EELTANYLLDHGHEFDQQQ 405

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2036	100.0	405	21	Maize Rad23 protei
2	1199.5	58.9	378	21	Arabidopsis thalia
3	1199.5	58.9	378	21	Arabidopsis thalia
4	1095	53.8	335	21	Arabidopsis thalia
5	1095	53.8	335	21	Arabidopsis thalia
6	981	48.2	307	21	Arabidopsis thalia
7	981	48.2	307	21	Arabidopsis thalia
8	945.5	46.4	371	21	Arabidopsis thalia
9	913	44.8	368	21	Arabidopsis thalia

10	900.5	44.2	368	21	AAV71459	Maize Rad23 protei
11	881.5	43.3	348	21	AAV71458	Arabidopsis thalia
12	850	41.7	345	21	AAV71457	Arabidopsis thalia
13	754.5	37.1	299	21	AAV71456	Arabidopsis thalia
14	732	36.0	257	21	AAV71455	Arabidopsis thalia
15	725	35.6	296	21	AAV71454	Arabidopsis thalia
16	634	31.1	409	19	AAV75700	Vpr protein bindin
17	634	31.1	409	19	AAV68186	Vpr protein bindin
18	632	31.0	409	24	ABU07460	Protein different
19	628.5	30.9	214	21	AAV71453	Arabidopsis thalia
20	623.5	30.6	416	23	ABV57171	Mouse ischaemic co
21	573	28.1	363	18	AAV23658	E6AP-binding prote
22	573	28.1	363	19	AAV75699	Vpr protein bindin
23	573	28.1	363	19	AAV68185	Vpr protein bindin
24	573	28.1	375	24	ABU11879	Human ABCAL intera
25	572.5	28.1	379	21	AAV58841	Breast and ovarian
26	548	26.9	346	18	AAV21730	Nuclear mitotic ap
27	514.5	25.3	186	21	AAV36531	Arabidopsis thalia
28	481	23.6	414	22	ABV58584	Drosophila melanog
29	347	17.0	179	19	AAV75843	Vpr protein bindin
30	347	17.0	179	19	AAV68187	Vpr protein bindin
31	324	15.9	174	19	AAV75850	Protein sequence o
32	324	15.9	174	19	AAV68194	Protein sequence o
33	318	15.6	117	23	ABP34618	Human ORF3591 prot
34	317	15.6	290	22	ABV72011	Drosophila melanog
35	307	15.1	136	21	AAV33942	Arabidopsis thalia
36	289	14.2	150	19	AAV75849	Protein sequence o
37	289	14.2	150	19	AAV68193	Protein sequence o
38	229	11.2	112	19	AAV75848	Protein sequence o
39	229	11.2	112	19	AAV68192	Protein sequence o
40	211	10.4	536	21	AAV30040	Arabidopsis thalia
41	211	10.4	577	21	AAV30039	Arabidopsis thalia
42	201	9.9	604	22	ABG16557	Novel human diagno
43	197.5	9.7	596	22	AAV64210	Murine HSP47 inter
44	197	9.7	589	22	AAV94008	Human stomach canc
45	197	9.7	589	22	AAV94311	Human protein sequ

ALIGNMENTS

RESULT 1

AAV71458
ID AAV71458 standard; Protein; 405 AA.

AC AAV71458;

DT 04-OCT-2000 (first entry)

DE Maize Rad23 protein #1.

XX Rad23; maize; ATCC No: PTA-530; recombinant expression cassette; wheat;
XX transgenic plant; soybean; sunflower; sorghum; canola; modulator.

OS Zea mays.

XX WO200031268-A1.

XX 02-JUN-2000.

XX 12-OCT-1999; 99WO-US24129.

XX 23-NOV-1998; 98US-0109728.

XX (PTON-) PIONEER HI-BRED INT INC.

XX Mahajan PB, Tagliani L;

XX WPI; 2000-400078/34.

XX N-PSDB; AAD01230.

PT Isolated nucleic acid encoding maize RAD23 protein is used to modulate
the levels of polypeptides in plant or in assays for identifying

PT compounds that bind to and/or increase/decrease enzymatic activity of
PT catalytically active polypeptides -
XX Claim 11b; Page 75-76; 82pp; English.
XX
CC The present sequence is the maize Rad23 protein #1. It is isolated from
CC V5 root tissue of a Zea mays cell line B73, infected with corn root
CC worm. The cDNA sequence is deposited under the ATCC No: PTA-530. Maize
CC Rad23 DNA sequence operably linked to a promoter can be used to construct
CC a recombinant expression cassette. This expression cassette can be used
CC to generate a dicot or monocot transgenic plant e.g., maize, soybean,
CC sunflower, sorghum, canola, wheat, etc.. It can also be used to modulate
CC the levels of Rad23 polypeptide expression in a plant or in assays to
CC identify compounds, that bind to and/or modulate the enzymatic activity
CC of catalytically active polypeptides.
XX
SQ Sequence 405 AA;

Query Match 100.0%; Score 2036; DB 21; Length 405;
Best Local Similarity 100.0%; Pred. No. 6.5e-138;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKNVTKLTGNTFEIRASPDASVADVKRIETTGQSTVRADQOMLIYQGIKILKDETTLLE 60
DB 1 MKNVTKLTGNTFEIRASPDASVADVKRIETTGQSTVRADQOMLIYQGIKILKDETTLLE 60
QY 61 SNGVAENSLVIMLSKAKSSGASTATTAKAPATLAQPAAPVAPASVARTPTQAPVAT 120
DB 61 SNGVAENSLVIMLSKAKSSGASTATTAKAPATLAQPAAPVAPASVARTPTQAPVAT 120
QY 121 AETAPSVQQAAPATVAATDDADVYSQAASNLVFGNNLEQTIQIILDMGGGTWERDVT 180
DB 121 AETAPSVQQAAPATVAATDDADVYSQAASNLVFGNNLEQTIQIILDMGGGTWERDVT 180
QY 181 VRALRAAYNNPERAIDYLSGIPENVEAQPVAPAAAGQOTNQQAASPAQPAVALPVQPS 240
DB 181 VRALRAAYNNPERAIDYLSGIPENVEAQPVAPAAAGQOTNQQAASPAQPAVALPVQPS 240
QY 241 PASAGPNANPLNFPQGVPSGGNPGVVPAGSGALDNLROLFOALLQLVQANPQILQ 300
DB 241 PASAGPNANPLNFPQGVPSGGNPGVVPAGSGALDNLROLFOALLQLVQANPQILQ 300
QY 301 PMLQELGKQNPQILRLIQENQAEFLRLVNESPEGGPGNLTGQLAAVPTLTVTPEERE 360
DB 301 PMLQELGKQNPQILRLIQENQAEFLRLVNESPEGGPGNLTGQLAAVPTLTVTPEERE 360
QY 361 AIORLEGWGNRELVLVEVFPACNKDELTANYLLDHGHEFDQOQ 405
DB 361 AIORLEGWGNRELVLVEVFPACNKDELTANYLLDHGHEFDQOQ 405

RESULT 2
AAG19974
ID AAG19974 standard; Protein; 378 AA.
XX
AC AAG19974;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 21981.
DE Arabidopsis thaliana
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX

25-FEB-1999; 99US-0121825.
05-MAR-1999; 99US-0123180.
09-MAR-1999; 99US-0123548.
23-MAR-1999; 99US-0125788.
25-MAR-1999; 99US-0126264.
29-MAR-1999; 99US-0126785.
01-APR-1999; 99US-0127462.
08-APR-1999; 99US-0128234.
16-APR-1999; 99US-0128714.
19-APR-1999; 99US-0129845.
21-APR-1999; 99US-0130077.
23-APR-1999; 99US-0130449.
28-APR-1999; 99US-0130510.
30-APR-1999; 99US-0130891.
04-MAY-1999; 99US-0131449.
06-MAY-1999; 99US-0132048.
07-MAY-1999; 99US-0132407.
11-MAY-1999; 99US-0132484.
14-MAY-1999; 99US-0132485.
14-MAY-1999; 99US-0132486.
14-MAY-1999; 99US-0132487.
14-MAY-1999; 99US-0132863.
14-MAY-1999; 99US-0134256.
14-MAY-1999; 99US-0134218.
14-MAY-1999; 99US-0134219.
14-MAY-1999; 99US-0134221.
14-MAY-1999; 99US-0134370.
18-MAY-1999; 99US-0134768.
19-MAY-1999; 99US-0134941.
20-MAY-1999; 99US-0135124.
21-MAY-1999; 99US-0135353.
24-MAY-1999; 99US-0135629.
25-MAY-1999; 99US-0136021.
27-MAY-1999; 99US-0136392.
28-MAY-1999; 99US-0136782.
01-JUN-1999; 99US-0137222.
03-JUN-1999; 99US-0137528.
04-JUN-1999; 99US-0137502.
07-JUN-1999; 99US-0137724.
08-JUN-1999; 99US-0138094.
10-JUN-1999; 99US-0138540.
10-JUN-1999; 99US-0138847.
14-JUN-1999; 99US-0139119.
16-JUN-1999; 99US-0139452.
16-JUN-1999; 99US-0139453.
17-JUN-1999; 99US-0139453.
18-JUN-1999; 99US-0139454.
18-JUN-1999; 99US-0139455.
18-JUN-1999; 99US-0139456.
18-JUN-1999; 99US-0139457.
18-JUN-1999; 99US-0139458.
18-JUN-1999; 99US-0139459.
18-JUN-1999; 99US-0139460.
18-JUN-1999; 99US-0139461.
18-JUN-1999; 99US-0139462.
18-JUN-1999; 99US-0139463.
18-JUN-1999; 99US-0139750.
18-JUN-1999; 99US-0139763.
21-JUN-1999; 99US-0139817.
22-JUN-1999; 99US-0139899.
23-JUN-1999; 99US-0140353.
23-JUN-1999; 99US-0140354.
24-JUN-1999; 99US-0140695.
28-JUN-1999; 99US-0140823.
29-JUN-1999; 99US-0140991.
30-JUN-1999; 99US-0141287.
01-JUL-1999; 99US-0141842.
01-JUL-1999; 99US-0142154.
02-JUL-1999; 99US-0142055.
06-JUL-1999; 99US-0142390.
08-JUL-1999; 99US-0142803.
09-JUL-1999; 99US-0142920.
12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145195.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.

PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 58.9%; Score 1199.5; DB 21; Length 378;
Best Local Similarity 62.9%; Pred. No. 5.8e-78;
Matches 256; Conservative 43; Mismatches 75; Indels 33; Gaps 9;

Qy 1 MCLNVKTLKGTNFEIEASPDASVADVKRIITETGQSTYRADQOQMLYQKILKDETTL 60
Db 1 MKIFVKTLSGSNFEIEVKPADKVDKTAIETVKG-AYEPAKQMLIHQKVLKDETTL 59
Qy 61 SNGVAENSFLVIMLSKAKASSSGASTAT---TAKAPATLAQPAAPVAPASVARTPTQA 116
Db 60 ENNVVENSFVIMLSKTKASPGASTASAPSPATQPTVATPOVS-APTASV----- 111
Qy 117 PVATAETAPPSVQQAAPAAATVAATDDADYVQAASNLVFGNNLEQTIQIILDMGGTWE 176
Db 112 PVPTSGTA-----TAAAPA--TAASVQTVYVQAASNLVAGTTLESTVQIILDMGGSWD 164
Qy 177 RDTVVRALRAAYNNPERAIDYLSGIPENVEAQVAPAPAAAGQOTNOQOASPAQPAVALP 236
Db 165 RDTVVRALRAAFNNPERAVEYLSGIPQAQAEIPPAQAQATGEOANPLAQPOQAAA--- 221
Qy 237 VQPSASAGNANPLNLFPOGVPSGSGNPGVPGAGSGALDRLQLPQFOLLQVQANP 296
Db 222 --PAATGGFNANPLNLFPOGMPAADA-----GAGAGNLDFLNSQFQALRAMVQANP 273
Qy 297 QILQPMLOELGKQNPQILRIQENQABFLRLVNESPEGGPGGNILGOLAAVPTLTVTTP 356
Db 274 QILQPMLOELGKQNPQLVRLIQEADFLRLINEPVEGES--NVMEQLEAAMPQAVTTP 331
Qy 357 EEREATORLEGMGFNRELVLVEFFACNCKDELTANYLLDHGHPDDQ 403
Db 332 EEREATERLEGMGFDRAWVLVEFFACNCKBELAANYLLDHMHFEFDQ 378

RESULT 3
AAG45204
ID AAG45204 standard; Protein; 378 AA.
XX
AC AAG45204;

XX 18-OCT-2000 (first entry)
DT Arabidopsis thaliana protein fragment SEQ ID NO: 56723.
XX
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
KW
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
XX
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137502.
PR 04-JUN-1999; 99US-0137724.
PR 07-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145085.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.


```
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158299.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160743.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160988.
PR 25-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 58.9%; Score 1199.5; DB 21; Length 378;
Best Local Similarity 62.9%; Pred. NO. 5.8e-78;
Matches 256; Conservative 43; Mismatches 75; Indels 33; Gaps 9;

Oy 1 MKNVKTLLKGNFRTFASPDASVADVKRIETTGOSTYRADQOMLIYQKILKDETTLE 60
Db 1 MKIFVKTUGSNFNELEVKPADKVSVDKTAIETVKG-AEYPAKQKLIHQGKVLKDETTLE 59

Oy 61 SNGVAENSFLVIMLSKAKASSGASTAT-----TAKAPATLAQPAAPVAPAPASVARTPTQA 116
Db 60 ENNVVENSFVIMLSKTKASPSGASTASAPAPSATQPTQVATPQVS-APTASV----- 111

Oy 117 PVATAETAPPVQQAAPAAATVAATDDADVYSQAASNLVFGNNLEQTIQQILDMGGGTWE 176
Db 112 FVPTSGTA-----TAAAPA--TAASVQTDVYQQAASNLVAGTTILESTVQQILDMGGGSWD 164

Oy 177 RDTVVRALAAVNNPERAIDLYLSGIPENVEAQPVARAPAGQQTNQQAASPAQPAVALP 236
Db 165 RDTVVRALAAVNNPERAVEYLYSGIPQAQAEIPVPAQAPATGEQAANPLAQPOQAAA--- 221

Oy 237 VQSPASAGPNANPLNLPQCGVPSGSGSNPGVVPVPGAGSGALDALRQLPQFQALLQLVQANP 296
```

```
Db 222 --PAAATGGPNANPLNLPQCGMPAADA-----GAGAGNLDFLNSQFQALRAMVQANP 273
Oy 297 QILOPMLQELGKQNPQILRIQENQAEFLRLVNESPGGPGGNTLGGAAAAVPQTLTTP 356
Db 274 QILOPMLQELGKQNPQILRIQEHQADFLRLINEPVEGER--NVMSQLEAAMPQAVTTP 331
Oy 357 EEREAIQRLEGMGFNRELVLVEFPACNKBELTANYLLDHGHEFDDQ 403
Db 332 EEREAIERLEGMGFDRAWVLVEFPACNKBELAAANYLLDHMHFEFDQ 378

RESULT 4
AAG19975
ID AAG19975 standard; Protein; 335 AA.
XX
AC AAG19975;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 21982.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
```

PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 08-AUG-1999; 99US-0147192.
PR 08-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157573.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 53.8%; Score 1095; DB 21; Length 335;
Best Local Similarity 63.9%; Pred. No. 1.6e-70;
Matches 232; Conservative 37; Mismatches 62; Indels 32; Gaps 8;

QY 45 MLIYQKILKDETTLESNGVAENSLVIMLSKAKASSSGASTAT-----TAKAPATLAQPA 100
DB 1 MLIHQKVLKDETTLEENNVVENSFIVIMLSKTKASPSSGASTASAPASATQPTVATPQ 60
QY 101 APVAPAAAVARTPTQAPVATAEATPPSVQPOQAAATAATDDADVYSQAASNLVFGNNL 160
DB 61 VS-APTASV-----PVPTSGTA-----TAAPA--TAASVQTDVYQAAASNLVAGTTL 105
QY 161 ECTIQIILDMGGTWERDVTVVRALRAAYNNPERAIDYLSGIPENVERAQPVARAPAAQOQ 220
DB 106 ESTVQIILDMGGSDRDTVVVRALRAAFNNPERAVEYLYSGIPAAAEIPPPVAQAPATGEQ 165
QY 221 TNQQAASPAQPAVALPVOPSPASAGPNANPLNLPQGVPSGSGNPGVVGAGSGALDALR 280
DB 166 AANPLAQPOQAAA-----PAAATGPNANPLNLPQGMFAADA-----GAGAGNLDFLR 214
QY 281 QLPOFQALLQVQANPQILOPMLQELGKONPQILRLQENQAEFLRLVNESPEGGPGNI 340
DB 215 NSQQFQARUAVQANPQILOPMLQELGKONPQILRLQENQAEFLRLVNESPEGGPGNI--NV 272
QY 341 LGQAAAAPQTLVTPPEREATQRLGNGFNRELVLVFFACNKDEELTANYLLDHGHEF 400
DB 273 MEQLEAANPQAVTVTPPEREATQRLGNGFNRELVLVFFACNKDEELTANYLLDHGHEF 332
QY 401 DDQ 403
DB 333 EDQ 335
RESULT 5
AAG45205
ID AAG45205 standard; Protein; 335 AA.
AC AAG45205;
XX
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 56724.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.

PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142055.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147152.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.

PR	12-OCT-1999;	99US-0158369.	OS	Arabidopsis thaliana.
PR	13-OCT-1999;	99US-0159293.	XX	
PR	13-OCT-1999;	99US-0159294.	PN	EP1033405-A2.
PR	13-OCT-1999;	99US-0159295.	XX	
PR	14-OCT-1999;	99US-0159329.	PD	06-SEP-2000.
PP	14-OCT-1999;	99US-0159330.	XX	
PR	14-OCT-1999;	99US-0159331.	PF	25-FEB-2000; 2000EP-0301439.
PR	14-OCT-1999;	99US-0159637.	XX	
PR	14-OCT-1999;	99US-0159638.	PR	25-FEB-1999; 99US-0121825.
PR	18-OCT-1999;	99US-0159584.	PR	05-MAR-1999; 99US-0123180.
PR	21-OCT-1999;	99US-0160741.	PR	09-MAR-1999; 99US-0123548.
PR	21-OCT-1999;	99US-0160767.	PR	23-MAR-1999; 99US-0125788.
PR	21-OCT-1999;	99US-0160768.	PR	25-MAR-1999; 99US-0126264.
PR	21-OCT-1999;	99US-0160770.	PR	29-MAR-1999; 99US-0126785.
PR	21-OCT-1999;	99US-0160814.	PR	01-APR-1999; 99US-0127462.
PR	21-OCT-1999;	99US-0160815.	PR	06-APR-1999; 99US-0128234.
PR	22-OCT-1999;	99US-0160980.	PR	08-APR-1999; 99US-0128714.
PR	22-OCT-1999;	99US-0160981.	PR	16-APR-1999; 99US-0129845.
PR	22-OCT-1999;	99US-0160989.	PR	19-APR-1999; 99US-0130077.
PR	25-OCT-1999;	99US-0161404.	PR	21-APR-1999; 99US-0130449.
PR	25-OCT-1999;	99US-0161405.	PR	23-APR-1999; 99US-0130510.
PR	25-OCT-1999;	99US-0161406.	PR	28-APR-1999; 99US-0130891.
PR	26-OCT-1999;	99US-0161359.	PR	30-APR-1999; 99US-0131449.
PR	26-OCT-1999;	99US-0161360.	PR	30-APR-1999; 99US-0132048.
PR	26-OCT-1999;	99US-0161361.	PR	04-MAY-1999; 99US-0132407.
PR	28-OCT-1999;	99US-0161920.	PR	04-MAY-1999; 99US-0132484.
PR	28-OCT-1999;	99US-0161992.	PR	05-MAY-1999; 99US-0132485.
PR	28-OCT-1999;	99US-0161993.	PR	06-MAY-1999; 99US-0132486.
PR	29-OCT-1999;	99US-0162142.	PR	07-MAY-1999; 99US-0132487.
Query Match 48.2%; Score 981; DB 21; Length 307;			PR	07-MAY-1999; 99US-0132863.
Best Local Similarity 62.7%; Pred. No. 2.le-62;			PR	11-MAY-1999; 99US-0134256.
Matches 210; Conservative 34; Mismatches 59; Indels 32; Gaps 8;			PR	14-MAY-1999; 99US-0134218.
			PR	14-MAY-1999; 99US-0134219.
			PR	14-MAY-1999; 99US-0134221.
QY	73	MLSKAKASSGASTAT-----TAKAPATLAQPAAPVAPAAASVARTPTQAPVATAETAPPSV 128	PR	14-MAY-1999; 99US-0134370.
Db	1	MLSKTASPSGASTASAPASATQPTVATPQVS-APTASV-----PVPTSGTA---- 48	PR	18-MAY-1999; 99US-0134768.
QY	129	QPOAAATVAATDDADVTSQASNLVFGNNLEQTTQQLDMGGTWERDVTVVRALRAAY 188	PR	19-MAY-1999; 99US-0134941.
Db	49	-TAAAPFA--TAASQVTDVYQGAASNLVAGTTLESTVQQLDMGGSWDRDVTVVRALRAAF 105	PR	20-MAY-1999; 99US-0135124.
QY	189	NNPERAIDLYSGIPENVEAQPVARAPAAQOQTNQQAASPAQPAVALPQPSPASAGPNA 248	PR	21-MAY-1999; 99US-0135353.
Db	106	NNPERAVEVLYSGIPAEIPFVQAAPATGEQANPLAQPQAAA-----PAAATGSPNA 160	PR	24-MAY-1999; 99US-0135629.
QY	249	NPLNLPQGVPSGGSPGVVPGAGSALDALROLQFQALLQLVQANPQILQPMQLQELGK 308	PR	25-MAY-1999; 99US-0136021.
Db	161	NPLNLPQGNPAADA-----GAGAGNLDFLRNSQQFQALRAMVQANPQILQPMQLQELGK 214	PR	27-MAY-1999; 99US-0136392.
QY	309	ONPQILRLIQENQAEFLRVNESPEGGPGGNILGOLAAAVPQTLTVPTEERAIQBLEGM 368	PR	01-JUN-1999; 99US-0137222.
Db	215	ONPOLVRLQEQHQADEFLRLINEVVEGEE--NVNMEQLAAAPQAVTVPTEERAIERLEGM 272	PR	03-JUN-1999; 99US-0137528.
QY	369	GFNRELVLVEFFACNCKDEELTANYLLDHGHEFDQ 403	PR	04-JUN-1999; 99US-0137502.
Db	273	GFDRAMVLEVFFACNKEELAANYLLDHHHEFDQ 307	PR	07-JUN-1999; 99US-0137724.
RESULT 7			PR	08-JUN-1999; 99US-0138094.
AAG45206			PR	10-JUN-1999; 99US-0138540.
ID	AAG45206 standard; Protein; 307 AA.		PR	10-JUN-1999; 99US-0138847.
XX	AAG45206;		PR	14-JUN-1999; 99US-0139119.
AC	18-OCT-2000 (first entry)		PR	16-JUN-1999; 99US-0139452.
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 56725.		PR	16-JUN-1999; 99US-0139453.
DT	Protein identification; signal transduction pathway; metabolic pathway;		PR	17-JUN-1999; 99US-0139492.
XX	hybridisation assay; Genetic mapping; gene expression control; promoter;		PR	18-JUN-1999; 99US-0139454.
DE	termination sequence.		PR	18-JUN-1999; 99US-0139455.
XX			PR	18-JUN-1999; 99US-0139456.
KW			PR	18-JUN-1999; 99US-0139457.
KW			PR	18-JUN-1999; 99US-0139458.
KW			PR	18-JUN-1999; 99US-0139460.
XX			PR	18-JUN-1999; 99US-0139461.
XX			PR	18-JUN-1999; 99US-0139462.
XX			PR	18-JUN-1999; 99US-0139463.
XX			PR	18-JUN-1999; 99US-0139750.
XX			PR	21-JUN-1999; 99US-0139763.
XX			PR	22-JUN-1999; 99US-0139817.
XX			PR	23-JUN-1999; 99US-0139899.
XX			PR	23-JUN-1999; 99US-0140353.
XX			PR	23-JUN-1999; 99US-0140354.
XX			PR	24-JUN-1999; 99US-0140695.
XX			PR	28-JUN-1999; 99US-0140823.
XX			PR	29-JUN-1999; 99US-0140991.

```

PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148664.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 48.2%; Score 981; DB 21; Length 307;
Best Local Similarity 62.7%; Pred. No. 2.1e-62;
Matches 210; Conservative 34; Mismatches 59; Indels 32; Gaps 8;

Qy 73 MLSKAKASSGASTAT----TAKAPATLAQPAAPVAPAAASVARTPTQAPVATATAPPSV 128
Db 1 MLSKTKASPSGASTASAPASATQPTVATPQVS-APTASV-----PVPTSCTA---- 48

Qy 129 QPQAPAPATVAATDDADVYSQAASNLVFGNNLEQTIOOILDMGGTWERDITVVRALRAAY 188
Db 49 -TAAAPA--TAASVQTDVYGOAASNLVAGTTLESTVQOILDMGGSGWDRTVVRAALRAAF 105

Qy 189 NNPERAIDLYSGIPENVEAQPVARAPAAQQTNOQAASPAQPAVALPVOPSPASAGPNA 248
Db 106 NNPERAVEYLYSGIPQAQAEIPVQAQAPATGEQAANPLAQOQAAA-----PAAATGGPNA 160

Qy 249 NPLNLPQGVPSGSGNPGVVGAGSGALDALRQLPQFQALLQLVQANPQILQPMQLBLGK 308
Db 161 NPLNLPQGMFAADA-----GAGAGNLDPLRNSQQFQALRAMVQANPQILQPMQLBLGK 214

Qy 309 QNPQILRLIQENQAEFLRLVNESPEGPGGNIIGOLAAAVPQILVTPPEREAIQRLEGM 368
Db 215 QNPQLVRLIOHQADFLRLINEPEVEGEE--NVMEQLEAAMPQAVTVTPPEREAIQRLEGM 272

Qy 369 GFNRRLVLEVFACNKDEELTANYLLDGHGFEDDQ 403
Db 273 GFDRAMVLEVFACNKNEELAAANYLLDGHMEFEDDQ 307

RESULT 8

```

AG17377
ID AG17377 standard; Protein; 371 AA.
XX AC RAG17377;
XX AC RAG17377;
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 18372.
DE Arabidopsis thaliana.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX EP1033405-A2.
PN 06-SEP-2000.
PD 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 07-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.


```
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154079.
PR 22-SEP-1999; 99US-0154739.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.

Query Match 46.4%; Score 945.5; DB 21; Length 371;
Best Local Similarity 50.7%; Pred. No. 9.3e-60;
Matches 208; Conservative 60; Mismatches 95; Indels 47; Gaps 11;

QY 1 MKNLVTKLGTNFEIEASPDASVADVKRIETTGQSTVRADQOQLIYQGIKDKETTL 60
DB 1 MKNLVTKLGSHEFIRVLPSDTIMAVKNIEDSQKDNYPCQQLIHNGKVLKDKETSLV 60

QY 61 SNGVAENFLVIMLSKAKASSGASTATTAKAPATLAQPAAPVAPAAASVARTPTQAPVAT 120
DB 61 ENKVTEEGFLVIMLSKSK---SGGSAGQASVQTSSVQPV-----SATTSSTKPAAP-ST 111

QY 121 AETAPSPVQP---QAAPATVAATDADVYSQAASNLVFGNNLEQTIQQILDMGGCTWER 177
DB 121 TOSSEVPASPAPQAPQAPAA-----QTDYTGQAASLVSGSLEQMQVQIMEMGGGWDK 165

QY 178 DTVVTRALRAAYNNPERAIDLYSGIPENVE-AQPVARAPAAQQTNQQAASPAQPAVALP 236
```

```
DB 166 ETVTRALRAAYNNPERAIDLYSGIPQTAEVAVPVEAQIAG-----SGAAP 212
QY 237 VQSPASAGNANPLNLFPGVPSGGSGNPGVVPAGS---GALDALRQLPQFALLQLVQ 293
DB 213 V--APASGGPNSSPLDLFPQ-----ETVAAAGSGDLGTLEFLRNDQFQQLRTMVH 261
QY 294 ANPQILQPMQLQELGKQNPQILRLIQENQABFLRLVNESPEGGPG-GNILGQAAAAVPQTL 352
DB 262 SNPQILQPMQLQELGKQNPQILRLIQENQABFLRLVNEPYEGSDGEGMDPQPEQEMPHAI 321
QY 353 TVTPEBEATQRLGKGFNRELVLVVFACNKBELTANYLLDHGHEPDD 402
DB 322 NVTPEBEATQRLGKGFNRELVLVVFACNKBELTANYLLDHGHEPDD 371

RESULT 9
AAG44341
ID AAG44341 standard; Protein; 368 AA.
XX
AC AAG44341;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55531.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
```

PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140895.
PR 24-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 44.8%; Score 913; DB 21; Length 368;
Best Local Similarity 47.5%; Pred. No. 2e-57;
Matches 194; Conservative 64; Mismatches 104; Indels 46; Gaps 7;

QY 1 MKNVKTLLKGTNFEIEASPDASADVKKRIETTCQSYRADQQMLIYQGIKLDKETTLE 60
DB 1 MKNVKTLLKGTNFEIEASPDASADVKKRIETTCQSYRADQQMLIYQGIKLDKETTLE 60
QY 61 SNGVAENSLVIMLSKAKASGASTATTAKAPATLAQPAAPVAPAAASVARTPTQAPVAT 120
DB 61 ENKVTEGFLVWLSKSTASSGPSST-----QPTS-----TTTSTISST 101
QY 121 AETAPPSVQPAAPAAVTA-----ATDDAVYSQAASHLVFGNNLEQTIQIILDMGGGTW 175
DB 102 TLAAPSTTQSIAPVAPASNSPTVQEQPTAQSDTYGQAASTLVSGSSTIEQMVQIIMEMGGGSM 161
QY 176 ERDVTVRALRAAYNNPERAIDVLYSGIPENVEAQVAPARAAGQQTNOQASPAQPAVAL 235
DB 162 DKETVTRALRAAYNNPERAIDVLYSGIPETV---TIPATNLSGVSGSRELTAP----- 211
QY 236 PVQSPASAGNANPLNLFPGVPSGGSNPGVVPQAGSGALDALRQLPQFQALLQVQAN 295
DB 212 -----PPSGGNSPLDLFPQEAUSDAA-----GGDLGTLEFLGNDOFQQLRSMVNSN 260
QY 296 POILQPMLOELGKQNPQILRLIQENQASFLRLVNESPPGGPGG-NILGQLAAAVPQTUTV 354
DB 261 POILQPMLOELGKQNPQILRLIQENQASFLRLVNESPPGGPGG-NILGQLAAAVPQTUTV 354
QY 355 TPREEAATQRLGCMGFNRELVLVEFPACNDELTANYLLDHGHEFDD 402
DB 321 TPEQESIERLEAMGFDRAIVIEAFLSCDRNEELAAANYLLEHSADFED 368

RESULT 10

AA771459
ID AA771459 standard; Protein; 368 AA.

AC AA771459;

DT 04-OCT-2000 (first entry)

DE Maize Rad23 protein #2.

KW Rad23; maize; ATCC No: PTA-531; recombinant expression cassette; wheat;
transgenic plant; soybean; sunflower; sorghum; canola; modulator.

OS Zea mays.

PN WO200031268-A1.

PD 02-JUN-2000.

PF 12-OCT-1999; 99WO-US24129.

PR 23-NOV-1998; 98US-0109728.

PA (PION-) PIONEER HI-BRED INT INC.

PI Mahajan PB, Tagliani L;

DR WPI; 2000-400078/34.

XX N-PSDB; AAD01231.

PT Isolated nucleic acid encoding maize RAD23 protein is used to modulate
the levels of polypeptides in plant or in assays for identifying
compounds that bind to and/or increase/decrease enzymatic activity of
catalytically active polypeptides -

PS Claim 11b; Page 78-79; 82pp; English.

CC The present sequence is the maize Rad23 protein #2. It is isolated from
a Zea mays cell line, B73 callus tissue regenerated five days after
transfer of the callus from medium containing auxin to a medium devoid

CC of exogenous auxin. The cDNA is deposited under the ATCC No: PTA-531.
CC Maize Rad23 DNA sequence operably linked to a promoter, can be used to
CC construct a recombinant expression cassette. This expression cassette
CC can be used to generate a dicot or monocot transgenic plant e.g., maize,
CC soybean, sunflower, sorghum, canola, wheat, etc. It can also be used to
CC modulate the levels of Rad23 polypeptide expression in a plant or in
CC assays to identify compounds, that bind to and/or modulate the enzymatic
CC activity of catalytically active polypeptides.

XX Sequence 368 AA;

Query Match 44.2%; Score 900.5; DB 21; Length 368;
Best Local Similarity 49.0%; Pred. No. 1.6e-56;
Matches 197; Conservative 53; Mismatches 117; Indels 35; Gaps 7;

QY 1 MKNVKTLLKGTNFEIEASPDASADVKKRIETTCQSYRADQQMLIYQGIKLDKETTLE 60

DB 1 MKNVKTLLKGTNFEIEASPDASADVKKRIETTCQSYRADQQMLIYQGIKLDKETTLE 60

QY 61 SNGVAENSLVIMLSKAKAS-SGASTATTAKAPATLAQPAAPVAPAAASVARTPTQAPVA 119

DB 61 ENKVNEGFLVWLSKGTSGTSSQHSNTPATRAP-----PLEAPQAP-QPPVA 114

QY 120 TAEAPPSVQPAAPAAVTAATDDAVYSQAASHLVFGNNLEQTIQIILDMGGGTWERDT 179

DB 115 PITTSQPEGLPAQAP-----NTHDNAASNLGSRNVDTIINOLMEMGGGSDKDK 164

QY 180 VVRALRAAYNNPERAIDVLYSGIPENVEAQVAPARAAGQQTNOQASPAQPAVALPQP 239

DB 165 VQRALRAAYNNPERAIDVLYSGIPETAB-----IAVIGGQANTTDRAPTGEA----- 213

QY 240 SPASAGNANPLNLFPGVPSGGSNPGVVPQAGSGALDALRQLPQFQALLQVQANPQIL 299

DB 214 -GLSGIPTAPDLFPQASNAGG-----GAGGGLDLFNNPFQAVREMVHTNPQIL 266

QY 300 QPMLOELGKQNPQILRLIQENQASFLRLVNESPPGGPGG-NILGQLAAAVPQTUTVPEER 359

DB 267 QPMVLVLSKQNPQILRLIENHDEFLQALNEPFGEGEDFLDQPEEDEMHAISVTPBEQ 326

QY 360 EAIQRLGCMGFNRELVLVEFPACNDELTANYLLDHGHEFD 401

DB 327 EAIQRLGCMGFNRELVLVEFPACNDELTANYLLDHGHEFD 368

RESULT 11

AA771378

ID AA771378 standard; Protein; 348 AA.

AC AA771378;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 18373.

KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.

PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.

PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 41.7%; Score 850; DB 21; Length 345;

Best Local Similarity 47.3%; Pred. No. 6e-53; Mismatches 60; Indels 46; Gaps 7;
Matches 181; Conservative 96;

Qy 26 VKRIETTGOGSTVRADQOMLIYOGKILKDETTLESNGVAENSLFVIMLSKAKASSSGAS 85
Db 3 VKKNIEDSQSKDNYPGQOLLIHNGKVLKDETTLVENKVTBEGFLVWMLSKSTASSAGP 62
Qy 86 TATTAKAPATLAQAPAPVAPAAASVARTPTQAPVATAETAPPSVQQAAPAAATVA-----A 140
Db 63 SST-----QPTS-----TTTSTISSTLAAPSTTQSIAPASNSTPVQEQP 103
Qy 141 TDDADVYQAAASNLVFGNNLEQTLQIILDMGGTWERDTVVRALRAAYNNPERAIDYLYS 200
Db 104 TAQSDTYGQAATLTVSGSSIEQMVOQIMEMGGGSKDKETVTRALRAAYNNPERAIDYLYS 163
Qy 201 GIPENVQAPVARAPAAAGQNTQQAASPAQPAVALPVQPSASAGPNANPLNLPQGVPS 260
Db 164 GIPETV---TIPATNLGSGVSGRELTAP-----PPSGGPNSSPLDLFQEAYS 208
Qy 261 GGSNPGVPGAGSGALDALRQLPQFALLQLVQANPQILQPMQLQBLGKQNPQILRLIQEN 320
Db 209 DAA-----GGDLGTLEFLRGNDQFQQLRSMVNSNPQILQPMQLQBLGKQNPQILRLIQEN 262
Qy 321 QAEFLRLVNESPEGGPGG-NILGQAAAAPQTLTTPPEREAIQRLGCHGFRNLELVF 379
Db 263 QAEFLQLLNEPYEGSDGVDIPDQDQEMPHSVNVTPEQESIERLEAMGFDRAIVIEAF 322

OY 380 FACNKDELTANYLLDHGHEPD 402
Db :.:.:| | | | | :.:|
323 LSCDRNEELANYLLEHSADFED 345

RESULT 13
AAG17379
ID AAG17379 standard; Protein; 299 AA.
XX AC AAG17379;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 18374.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0125264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145132.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 17-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 06:20:11 ; Search time 138.842 Seconds
(without alignments)
544.771 Million cell updates/sec

Title: US-09-805-550-2

Perfect score: 2036

Sequence: 1 MKNLVKTLKGTNFIEASPD.....BELTANYLLDRHGHEFDQQQ 405

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 696363 seqs, 186758610 residues

Total number of hits satisfying chosen parameters: 696363

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2036	100.0	405	9	US-09-805-550-2
2	900.5	44.2	368	9	US-09-805-550-2
3	573	28.1	363	12	US-10-116-275-170
4	572.5	28.1	379	15	US-10-102-806-549
5	198	9.7	595	12	US-10-293-000-2
6	197	9.7	589	12	US-10-293-000-4
7	166.5	8.2	624	15	US-10-146-473-47
8	151	7.4	79	11	US-09-918-036-4
9	149	7.3	1023	11	US-09-893-519A-14
10	146	7.2	681	9	US-09-815-242-11830
11	144	7.1	727	15	US-10-234-432-30
12	144	7.1	727	15	US-10-234-432-57
13	141.5	6.9	638	15	US-10-234-432-59
14	141.5	6.9	2971	15	US-10-146-473-50
15	140	6.9	2441	14	US-10-109-886-8

16	138.5	6.8	522	15	US-10-296-770-2	Sequence 2, Appli
17	138	6.8	81	11	US-09-918-036-5	Sequence 5, Appli
18	137	6.7	578	12	US-10-032-585-7793	Sequence 7793, Ap
19	137	6.7	810	15	US-10-156-761-10081	Sequence 10081, A
20	136	6.7	655	14	US-10-001-632A-2	Sequence 2, Appli
21	136	6.7	655	15	US-10-054-683-29	Sequence 29, Appli
22	135	6.6	352	11	US-09-820-843A-23	Sequence 23, Appli
23	134.5	6.6	699	12	US-10-200-562-143	Sequence 143, App
24	134.5	6.6	699	12	US-10-237-551-143	Sequence 143, App
25	134.5	6.6	699	12	US-10-237-551-254	Sequence 254, App
26	134.5	6.6	699	15	US-10-121-988-143	Sequence 143, App
27	134.5	6.6	1276	15	US-10-156-761-10509	Sequence 10509, A
28	134	6.6	1004	10	US-09-738-626-5676	Sequence 5676, Ap
29	133.5	6.6	358	12	US-10-024-298A-137	Sequence 137, App
30	133.5	6.6	358	12	US-10-042-211A-137	Sequence 137, App
31	133.5	6.6	358	15	US-10-197-666A-32	Sequence 32, Appli
32	133.5	6.6	396	12	US-10-024-298A-139	Sequence 139, App
33	133.5	6.6	396	12	US-10-042-211A-139	Sequence 139, App
34	133.5	6.6	396	15	US-10-197-666A-34	Sequence 34, Appli
35	133.5	6.6	718	12	US-10-029-386-34049	Sequence 34049, A
36	133.5	6.6	2442	14	US-10-109-886-10	Sequence 10, Appli
37	132.5	6.5	309	11	US-09-820-843A-24	Sequence 24, Appli
38	132	6.5	1259	15	US-10-260-715-8	Sequence 8, Appli
39	131.5	6.5	1744	14	US-10-108-605-25	Sequence 25, Appli
40	131	6.4	1247	15	US-10-128-714-3473	Sequence 3473, Ap
41	131	6.4	1263	15	US-10-128-714-8473	Sequence 8473, Ap
42	130	6.4	2382	15	US-10-196-935A-2	Sequence 2, Appli
43	129	6.3	63	9	US-09-925-299-1544	Sequence 1544, Ap
44	129	6.3	63	11	US-09-925-299-1544	Sequence 1544, Ap
45	129	6.3	504	11	US-09-938-864-410	Sequence 410, App

ALIGNMENTS

RESULT 1

US-09-805-550-2
; Sequence 2, Application US/09805550
; Patent No. US20020026045A1
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; APPLICANT: Tagliani, Laura
; TITLE OF INVENTION: Rad23 Genes and Uses Thereof
; FILE REFERENCE: 0964D
; CURRENT APPLICATION NUMBER: US/09/805.550
; CURRENT FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 09/413,574
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: 60/109,728
; PRIOR FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Zea mays
; US-09-805-550-2

Query Match 100.0%; Score 2036; DB 9; Length 405;
Best Local Similarity 100.0%; Pred. No. 2.3e-146;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MKNLVKTLKGTNFIEASPDASVADVKRIIETTCOSTYRADQOMLIYQKILKDETTLE	60
Db	1	MKNLVKTLKGTNFIEASPDASVADVKRIIETTCOSTYRADQOMLIYQKILKDETTLE	60
Qy	61	SNVAENSLVIMLSKAKASSGASTATTAKAPATLAQAPAPAPAPAPASVARTPTQAPVAT	120
Db	61	SNVAENSLVIMLSKAKASSGASTATTAKAPATLAQAPAPAPAPAPASVARTPTQAPVAT	120
Qy	121	AETAPPSVQQAAPAAVATDDADVVYSQAASNLVFGNNLEQTIOQILDMGGGTWERDVT	180
Db	121	AETAPPSVQQAAPAAVATDDADVVYSQAASNLVFGNNLEQTIOQILDMGGGTWERDVT	180

Qy 181 VRLRAAYNNPERAIDYLSGIPENVEAQVAPAPAGQQTNOQAASPAQVALPVQPS 240
Db 181 VRLRAAYNNPERAIDYLSGIPENVEAQVAPAPAGQQTNOQAASPAQVALPVQPS 240
Qy 241 PASAGPNANPLNFPQGVPSGSGNPGVPGAGSGALDRLQLPQFQALLQVQANPQILQ 300
Db 241 PASAGPNANPLNFPQGVPSGSGNPGVPGAGSGALDRLQLPQFQALLQVQANPQILQ 300
Qy 301 PMLQELGKONPQILRLIQENQAEFLRLVNESPEGGPGNLLGOLAAAVPOTLVTTPERE 360
Db 301 PMLQELGKONPQILRLIQENQAEFLRLVNESPEGGPGNLLGOLAAAVPOTLVTTPERE 360
Qy 361 AIQRLGGMGNRELVLVFFACNKDELTANYLLDHGHEFDQ 405
Db 361 AIQRLGGMGNRELVLVFFACNKDELTANYLLDHGHEFDQ 405

RESULT 2
US-09-805-550-4
; Sequence 4, Application US/09805550
; Patent No. US2002026045A1
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Rad23 Genes and Uses Thereof
; FILE REFERENCE: 0964D
; CURRENT APPLICATION NUMBER: US/09/805,550
; CURRENT FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 09/413,574
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: 60/109,728
; PRIOR FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Zea mays
US-09-805-550-4

Query Match 44.2%; Score 900.5; DB 9; Length 368;
Best Local Similarity 49.0%; Pred. No. 2.8e+60;
Matches 197; Conservative 53; Mismatches 117; Indels 35; Gaps 7;

Qy 1 MKNVTLKGTNFEIAPSDASVADVKRIETTTQGSTYRADQQMLIYQKILKDETTLE 60
Db 1 MKLTVTKLKGTHFEIRVQPNNDTINAVKKNIEIQKDSYPMGQQLLIFNGKVLKDESTLE 60
Qy 61 SNGVAENSLVIMLSKAKAS--SSGASTATTAKAPATLAQAPAPVAPAAASVARTPTQAPVA 119
Db 61 ENKYNEDGFLVVLMSKGTSGSTGTSSQHSNTPTATQAP-----PLEAPQAP-QPPVA 114
Qy 120 TAETAPSVQQAAPAAATDADVYSQAASNLVFGNNLEQITIQIILDMGGGTWERDT 179
Db 115 PITTSQPEGLPAQAP-----NTHDNAASLLSGRNVDTIINLMEMGGGSKDK 164
Qy 180 VRLRAAYNNPERAIDYLSGIPENVEAQVAPAPAGQQTNOQAASPAQVALPVQ 239
Db 165 VQRLAAYNNPERAVEYLYSGIPVTAE-----IAVPIGGQGANTTDRAPTGEA----- 213
Qy 240 SPASAGPNANPLNFPQGVPSGSGNPGVPGAGSGALDRLQLPQFQALLQVQANPQIL 299
Db 214 -GLSGIPNTAPDLDFPQASNAGG-----GAGGGPLDFLRNNPQFQAVREMTNPNQIL 266
Qy 300 QPMQLQELGKONPQILRLIQENQAEFLRLVNESPEGGPGNLLGOLAAAVPOTLVTTPERE 359
Db 267 QPMLVELSKONPQILRLIEENHDFLQNLNPEFEGEGDFLDQPEDEMHPAISVTPEEQ 326
Qy 360 EAIQRLGGMGNRELVLVFFACNKDELTANYLLDHGHEFD 401
Db 327 EAIQRLGSMGDFRVARVIEAFLACDRNEELAAANYLLEHAGEED 368

RESULT 3
US-10-116-275-170
; Sequence 170, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elian Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods an
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 170
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-275-170

Query Match 28.1%; Score 573; DB 12; Length 363;
Best Local Similarity 33.3%; Pred. No. 2e+35;
Matches 139; Conservative 79; Mismatches 124; Indels 76; Gaps 13;

Qy 3 LNVKTLKGTNFEIAPSDASVADVKRIETTTQGSTYRADQQMLIYQKILKDETTLESN 62
Db 5 ITLKTLQQOTFKIRMEPDETIVKLEKIEAKGRDAPPVAGQKLIYAGKILSDDVPIRDY 64
Qy 63 GVAENSLVIMLSKAKAS--SSGASTATTAKAPATLAQAPAA--PVAPAAASVARTPTQAPVA 119
Db 65 RIDERKFNVMVMTKKA---GQTSAPPEASPTAPESSTSPPTSGMSHPP---PAA 118
Qy 120 TAETAP-----PSVQQAAPAAATDADVYSQAASNLVFGNNLEQITIQIILDMGGGT 174
Db 119 REDKSPSEASPTTSPESV-SGSVPSSGSGREEDAASTLVTSGETVMTLIMSNG--- 174
Qy 175 WERTTVVRLAAYNNPERAIDYLSGIPENVEAQVAPAPAGQ-QTNQQAASPAQPAV 233
Db 175 YERERVVAALRASYNPHRAVEYLLTGIPGSP-----PEHGSVQESQVSEQPATEA- 226
Qy 234 ALPVQSPASAGPNANPLNFPQGVPSGSGNPGVPGAGSGALDRLQLPQFQALLQVQ 293
Db 227 -----AGENPLEFLERDQFPQNMQRQVIQ 249
Qy 294 ANPQILQPMQLQELGKONPQILRLIQENQAEFLRLVNESPEGGPGN-----LGQLAAA 347
Db 250 QNPALLPALLQQLGOENPOLLQOISRHOQFIQMLNEPP--GELADISDVEGEVGAIGEE 307
Qy 348 VPQT--LTVTPEREAIQRLGGMGNRELVLVFFACNKDELTANYLLDHGHEFDQ 403
Db 308 APQWNYIQTVPQEKAIERLKGALFPPESLVIOAYFACEKNELAAANPLL--SQNFDD 363

RESULT 4
US-10-102-806-549
; Sequence 549, Application US/10102806
; Publication No. US2003005421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270


```
Db 153 GGLAGLSSGLNTNFSELQSQ-----MORQLLSNPEMMYQIMENPFVQSMLSNPDLMRQ 207
QY 154 LVFGN-NLEQTIQILDMGGTWERDVTVVRALRAAYN-----NPERADLYSGIP 203
Db 208 LIMANFQMOQLIQRNPEISHMLNPNPDMIRQTLERARNPAMQEMMRNQDRLALS-IP 266
QY 204 ENVEA-----QFVARA-----PAAGQOTNQQAASPAQPAVALFVOPSPASAGP 246
Db 267 GGYNALLRRMYTDIQEPMLSAQAQFGNPFASLVNTSSGEGSQPSRTENRDLNPNWAP 326
QY 247 NANPLNLPFGVPS-----GGSN-----PGVPGAGSGALDALRQLPQFOALLQ 290
Db 327 QTSQSSASSTASTVGGTTGASTGTSQSTTAPNLVPGVGASMENT-----PGMQSLIQ 382
QY 291 LVQANPQILQ-----PMLQELGKQ----- 309
Db 383 QITENPQLMQLMSAPYMRSMQSLSONPDLAQAQMLNPNLPFAGNPQLQEQMRQQLPTFL 442
QY 310 -----NPOILRLIOENQAEFLRLVNESP-----EGPGG----- 338
Db 443 QMQNPDTLTSAQMSNPRAMQALLQIQOGLQTLATEAPGLIPGFTPGGLGALGSTGSSGTNG 502
QY 339 -----NILGOLAAAVPOTLTVTPEEREAIQRLGEMGF-NR 372
Db 503 SNATPSNTSPTAGTTPGHHQFIQMLQALAGVNPQLQNPVRFQOQLEQLSAMGFLNR 562
QY 373 ELVLEVFACNKEELTANVLL 394
Db 563 EANLQALATGDDINAAIERLL 584

RESULT 7
US-10-146-473-47
; Sequence 47, Application US/10146473
; Publication No. US20030108888A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseung
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 47
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-473-47
```

```
Query Match 8.2%; Score 166.5; DB 15; Length 624;
Best Local Similarity 22.2%; Pred. No. 0.00028;
Matches 100; Conservative 64; Mismatches 178; Indels 109; Gaps 20;
QY 1 MKNLVKTLKGTNFEIASPDASVADVVKRIETTGQSTYRADQOQLIYQKILKDETTLE 60
Db 33 IKVTPTPKKE-EFAPVENSNSVQFKEALSKFKSGT-----DQLVLIFAGKILKQDTLI 89
QY 61 SNGVAENSLVIMLSKAKASSGASTATTAKAPATLQAAPVAPAASVARTP--TQAPV 118
Db 89 QHGI-HDGLTVHLVIKSNRPQGST-----QPSNAAGTWTTSASTRSNSTPI 136
QY 119 AT-----AETAPPSVQFQ-----AAPAATVAATDDADVTYSQAA 151
Db 137 STNSNPFGLSLGLAGLSLGLSTNFSLSQMQOQLMASPEMMIQIMENPFVQSMLS 196
```

```
QY 152 S-----NLVFGN-NLEQTIQ-----QILDMGGTWERDVTVVRALRAAYN 189
Db 197 NPDLMRQLIMANPQMOQLIQRNPEISHMLNPNPDMIRQTLERARNPAMQEMMRNQDRLALS 256
QY 190 NPE-----RAIDYLYSGIPE-----NVEAQFVARAPAAGQOTNQQAASPAQPA-----V 233
Db 257 NLESTPGGYNALLRRMYTDIQEPMLNAAQEQFGNPFASVSSSSSGEGTQPSRTENRDL 316
QY 234 ALPVOPSPASAPNPNLNPFGQVPSGSGSNPGVVFG--AGSGALDALRQLPQFOALLQ 291
Db 317 PNPWAPPATQS-SATTTSTTSGSGSNSSNATGNTVAAANYVASIFSTPGMSLLQ 375
QY 292 VOANPQILQPM-----QELGKQNPQILRLIOENQAEFLRLVNESPEGPGGNILG 342
Db 376 ITENPQLIQNMLSAFYMRSMQSL--SONPDLAQAQMLNPLTANPQLQEQMRP----- 428
QY 343 QLAHAV-----PQTLTV--TPEEREAIQRL 366
Db 429 QLPFAFLOQMNPDTLTSAQMSNPRAMQALMIQ 459

RESULT 8
US-09-918-036-4
; Sequence 4, Application US/09918036
; Publication No. US20030092159A1
; GENERAL INFORMATION:
; APPLICANT: MADURA, Kitan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE RAPID PURIFICATION OF PROTEAS
; FILE REFERENCE: 266/165
; CURRENT APPLICATION NUMBER: US/09/918,036
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/050,171
; PRIOR FILING DATE: 1997-06-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-918-036-4

Query Match 7.4%; Score 151; DB 11; Length 79;
Best Local Similarity 38.0%; Pred. No. 0.00026;
Matches 30; Conservative 21; Mismatches 28; Indels 0; Gaps 0;
QY 1 MKNLVKTLKGTNFEIASPDASVADVVKRIETTGQSTYRADQOQLIYQKILKDETTLE 60
Db 1 MOVTLKTLQOQTFKIDIDPEETVKALKEIESEKGDAPPVAGOKLIYAGKILNDDTALK 60
QY 61 SNGVAENSLVIMLSKAKA 79
Db 61 EYKIDKKNFVVVVVTKPKA 79

RESULT 9
US-09-893-519A-14
; Sequence 14, Application US/09893519A
; Publication No. US20030027243A1
; GENERAL INFORMATION:
; APPLICANT: ANADYS PHARMACEUTICALS, INC.
; APPLICANT: THOMPSON, Craig
; APPLICANT: MOORE, Jeffrey
; APPLICANT: BUURMAN, Ed T.
; APPLICANT: BRADLEY, John
; APPLICANT: DESILVA, Thamara
; APPLICANT: HARRIS, Sandra
; APPLICANT: KOMARNITSKY, Svetlana
; APPLICANT: MENDELLO, Marc
; APPLICANT: MOORE, Daniel
; APPLICANT: MCCOY, Melissa
; APPLICANT: SANDERSON, Karen
; APPLICANT: HAQ, Tariq
```

; APPLICANT: ZHU, Shuhao
; APPLICANT: LONG, Fan
; APPLICANT: DAVIDOV, Eugene
; TITLE OF INVENTION: ANTIFUNGAL COMPOUNDS AND METHODS OF USE
; FILE REFERENCE: 0342/1G548-US2
; CURRENT APPLICATION NUMBER: US/09/893,519A
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,164
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 60/224,457
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 146
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 1023
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Corresponds to SEQ ID NO: 87
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Human Genbank/CAA72189
; DATABASE ENTRY DATE: 1997-06-25
; RELEVANT RESIDUES: (1)..(1023)
US-09-893-519A-14

Query Match 7.3%; Score 149; DB 11; Length 1023;
Best Local Similarity 26.1%; Pred. No. 0.012;
Matches 86; Conservative 31; Mismatches 128; Indels 84; Gaps 16;

QY 58 TLENGVAENSLVIMLSKAKSSGASTATT-----AKAPATLAQPAAPVAPA-ASVART 112
DB 214 SLVNGPA-----ALLPLKPAAPGTVIQTFFVGAAPAPAPAPAPSPAPAPAPAAA 267
QY 113 PTQAPVATETAPSVQQAAPATVAATDDADYVSOAASNLVFGNNLEQTIQOILDMGG 172
DB 268 PPPPPAPATLARPPGHPAGPPTAPAVPPPA-----AAQN-----G 304
QY 173 GTWERDVTVRALRAAYNNPERAIDLYSGIPENVEAQP-----VARAPAAAGOOTNQ----- 223
DB 305 GS-----AGAPAPAPA-----AGGPAAGVSGQPGPGAAAPAPAGVKAESPKEV 349
QY 224 QAASPAQPAVALPVQPSASA-----GPNANPLNLPFGV-----PSGSGNPGVVPFGAGSALDA 278
DB 350 QAAPPAQTLAASGASTAASVIGTQWQALPSPAAPPVPPAPGTPTGLPKGAAGAVTQS 409
QY 279 LRQLPQ-----FOALLQVLQANPQILQPMLOELKQNPQI-----LRLIQENQAEFLRLVNE 330
DB 410 LSRTPTATTSGTRATLTPTVLAPRLPQPPQNPNTNQFQLPPGMVLRSENGOLLMTIPQ 469
QY 331 SPEGGPGNIGQLAA-----AVPQTLTVTP 356
DB 470 A-----LAQMOQAHAQFPQT-TMAP 488

RESULT 10
US-09-815-242-11830
; Sequence 11830, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11830
; LENGTH: 681
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-11830

Query Match 7.2%; Score 146; DB 9; Length 681;
Best Local Similarity 23.7%; Pred. No. 0.011;
Matches 98; Conservative 56; Mismatches 164; Indels 96; Gaps 18;

QY 16 EASPDAAA---SVADVKKRIETTGQSTVRAQQQMLIVQKILKDETTLESNGVAENSLVI 72
DB 296 QALPEALDNGQGBRERVLAQAALPA---EDVQFYQWGLIGRRDLPLAPD---PRSGFEMV 351
QY 73 ML-----SKAKASSGASTATTAKAPATLAQPAAPVAPAASVARTPTQAPVA 119
DB 352 LLRLMAFPADAGVPRPTPLKLGISKATTDTPANSPVAGAAASP-APVATVAPVPAAPV 410
QY 120 TATAPSVQPAAPATVAATDDADYVSOAASNLVFGNNLEQTIQOILDMGGTWER-- 177
DB 411 EAPAAPPAAP-PSAPPAAPAAVEARVTEAVVVEEPAAA-----AEVVDL-----PWSEPA 454
QY 178 -----DVTVRALRAAYNNPERAIDLYSGIPENVEAQPVARAPAAAGOOTNQAAA 226
DB 455 PSLSAEPPEPEPEPEPLAVEAPSPVPAVAVAVETV---LEALPAALPVAPDEQDEQDE 511
QY 227 SP-----AQPVALPVQPSAPA-----SAGPNANPLNLPFQ-----GVPS 260
DB 512 PPPADDYVEVDMDTLAYLDATPEPDVVVVEEPLPAAPKATGLAAEWLELPRGLGLGTA 571
QY 261 G-GSNPGV-----PGAGSALDALQOLPOFQALLQVLQANPQILQPMLOELG 307
DB 572 SIGANCTVLAADDHWHHLHLDPGQ-SALFNATQORRLNDALNQHGLRTLK-LEVTLQKPE 629
QY 308 KONPQILRLIQENQAEFLRLVNESPEGGPGNIGQLAAAVPOTTLTVTPBEERA 361
DB 630 QETP--AQAARRRAERQRAAEASIDADPLVRQLRQEPFAAVVRDGTIEPLEAKA 681

RESULT 11
US-10-234-432-30
; Sequence 30, Application US/10234432
; Publication No. US20030091598A1
; GENERAL INFORMATION:
; APPLICANT: Homer, Mary J.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Persing, David H.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF BABESIA INFECTION
; FILE REFERENCE: 210121.560
; CURRENT APPLICATION NUMBER: US/10/234,432
; CURRENT FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 30


```
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Babesia sp. WA1
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 264
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-234-432-30

Query Match          7.1%; Score 144; DB 15; Length 727;
Best Local Similarity 24.4%; Pred. No. 0.018;
Matches 95; Conservative 45; Mismatches 169; Indels 80; Gaps 16;

Qy 17 ASDASVADVKRIETTCQSTYRADQOQMLIYQKILKDETTLESNGVAENSLVIMLSK 76
Db 72 ATPEASGSANQ-----QQSQTGAGESQPVLSSTEMATVKEETVPETK--VENNVV-----120

Qy 77 AKASSSGASTATTAKAPA--TLAQ-PAAPVAPAAASVARTPTQ-APVATAETAPPSVQPOA 132
Db 121 -----QEATVTPAQVPAVENVSQPTTQTVAAPAAPAQPOQPAQVAPQATAGIQQAQPQVA 174

Qy 133 APAAT-----VAATDDADVISOAASN-----LVFGNNLEOTIQIILDMGGGTWER 177
Db 175 TETATAEQPVAATTTTEVQMPQAAASPAPILETPOVMTQTAPVEETQAPVVTESPAPOQP 234

Qy 178 DTVVRALRAAYNNPERAIDYLYSGIPENVEAQPVARAQAAGQQTNQQAASPAQPAVALPV 237
Db 235 AQV-----AAPEQPAEVAPOATAGI-QQAQOPQVAVATETATAEQPVAAATTEVQMPQAAAE 288

Qy 238 QPSPASAGPNA-----NPLNLFPGQVPSGSGSNPGVVPAGSGALDALR 280
Db 289 SPAPISSETPOVMTQTAPVEETQAPVVTESPAPOQPAQVAAPEQPAEVAPOATAGIQ- --- 344

Qy 281 QLPQFOALLQVQANPOILOPMLQELGKQ---NPQILRLIQENQAEFLRLVNESPEGPG 337
Db 345 -----QAQOPQVAAEAQVQPPVQTAQTRPVAQPOV--VVAEAQ-----VVQPPVKAQA 391

Qy 338 GNIILQLAAAVPQTLTVPREAIQRL 366
Db 392 AQPVVKDQAQP-VASVAPQATAGIQQAQ 419

RESULT 13
US-10-234-432-59
; Sequence 59, Application US/10234432
; Publication No. US20030091598A1
; GENERAL INFORMATION:
; APPLICANT: Homer, Mary J.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Persing, David H.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND
; FILE REFERENCE: 210121.560
; CURRENT APPLICATION NUMBER: US/10/234, 432
; CURRENT FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 59
; LENGTH: 638
; TYPE: PRT
; ORGANISM: Babesia sp. WA1
US-10-234-432-59

Query Match          6.9%; Score 141.5; DB 15; Length 638;
Best Local Similarity 24.2%; Pred. No. 0.023;
Matches 89; Conservative 39; Mismatches 119; Indels 121; Gaps 17;

Qy 17 ASDASVADVKRIETTCQSTYRADQOQMLIYQKILKDETTLESNGVAENSLVIMLSK 76
Db 72 ATPEASGSANQ-----QQSQTGAGESQPVLSSTEMATVKEETVPETK--VENNVV-----120

Qy 77 AKASSSGASTATTAKAPA--TLAQ-PAAPVAPAAASVARTPTQ-APVATAETAPPSVQPOA 132
Db 121 -----QEATVTPAQVPAVENVSQPTTQTVAAPAAPAQPOQPAQVAPQATAGIQ- ---QAQP 170

Qy 133 APAATVAATDDADVISOAASNLVFGNNLEQTTQIILDMGGGTWERDITVVRALRAAYNNPE 192
Db 171 QPVATETAEQPV--AATT-----TEVQMPQAAAESPA 202

Qy 193 RAID-----YLYSGIPENVEAQPVARAQAAGQQTNQQAASPAQPAVALP-----VQPS 240
Db 203 PILETPOVMTQTAPVEETQAPVVTESPAPOQPA--QVAAPEQPAEVAPOATAGIQQAQPQ 260

Qy 241 PASA-----GPNANPLNLFPG-----VPSGSGSNP-----GVVP 269
Db 261 PVAAEAQVQPPVQTAQTRPVAQPOQVVAEAQVQPPVKAQAQPVVKDQAAPVASVAP 320

Qy 270 GAGSGALDALRQLPQFALLQLVQANPQILQPM-----QELGKQNPQILR 315
```

Db 321 QATAGIQ-----QAPQPVAAEAQVPPVKAARKPKPIVKDQAAQPVAPQATA 371
Qy 316 LIQENQAE 323
Db 372 GVAEDQSE 379

RESULT 14
US-10-146-473-50
; Sequence 50, Application US/10146473
; Publication No. US2003010888A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
; LENGTH: 2971
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-473-50

Query Match 6.9%; Score 141.5; DB 15; Length 2971;
Best Local Similarity 25.5%; Pred. No. 0.18;
Matches 95; Conservative 25; Mismatches 142; Indels 111; Gaps 16;
Qy 79 ASSGASTATTAKAPAT--LAQP-AAPV-----APAASVARTPTQAPVATAETAPPSVOP 130
Db 1177 SSSLPISVPTTLPAASAPLTIPISAPLTVSASGPAULTSVTPPLAPVVPAPGPPSLQP 1236
Qy 131 QAAPAAATVAATDDADVISOAASNLVFGNN--LEQTIOQILDGCGGTWERTVVRALRAAY 188
Db 1237 SCASPSASALTGLATAPSLSSQTPGHPLLLATSSHVGL-----NSTVAPACSPVL 1290
Qy 189 NNPERAIDYLGIGIPENVEAQVARAPA-----AGOOTNQQAASPAQPAVA 234
Db 1291 -VPASALASPPPSAPNPAPQAASLLAPASASQALATPLAPMAAPQTAILAPSPAPPLAP 1349
Qy 235 LPV-QPSPASA-----GPNANP-----LNLFPQGV 258
Db 1350 LPVLAPSGAAPVLASSQTPVVPMAPSPTGTSLASASPVPAPTVLPAPSTQTMLPAPV 1409
Qy 259 PS-----GGSNPGVVPVGGAGDALRLQLPQFQALLQLVQANPQ 297
Db 1410 PSLPSPASTQTALAPALAPTLTGSSPSQTLSTGN-----PQGFPTQTLSTLPASSL 1465
Qy 298 ILQP-MLQELGKNP-----QILRLIQENQAEFLRLVNESPEG-----GPGNIIQGO 343
Db 1466 VTPAQTLTSLAPGPPGLPTQLSL-----APAPPLAPASVPGPAPAHTLTLAPASSASL 1520
Qy 344 LAAAVPQTILTVP 356
Db 1521 LAPASVQTLTUSP 1533

RESULT 15
US-10-109-886-8
; Sequence 8, Application US/10109886
; Publication No. US20020119499A1
; GENERAL INFORMATION:
; APPLICANT: TANIGUCHI, TOMOYASU

; APPLICANT: MIZUKAMI, Junko
; TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND
; FILE REFERENCE: TANIGUCHI-6
; CURRENT APPLICATION NUMBER: US/10/109,886
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/514,247
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: PCT/JP98/03734
; PRIOR FILING DATE: 1998-08-24
; PRIOR APPLICATION NUMBER: JP231084/1997
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 2441
; TYPE: PRT
; ORGANISM: mouse
US-10-109-886-8

Query Match 6.9%; Score 140; DB 14; Length 2441;
Best Local Similarity 21.7%; Pred. No. 0.18; Mismatches 41; Indels 132; Gaps 18;
Matches 92; Conservative 41; Mismatches 159; Indels 132; Gaps 18;
Qy 44 QMLYQGKILKDE--TTLESNGVAENSLVIMLSKAKASSGASTATTAKAPATILAQPAAP 102
Db 1857 QHCLQQAQLMERRMATNTRNPQSL-----PSPTSAPFGTPTQOESTPTQTPPP 1907
Qy 103 VAPAS-----VART--PT-----QAPVATATAPSPVQPOAPAAATVAATDDAD 145
Db 1908 AQOPSPVMSPPAGFPNVARTQPTIVSAGKTNQVPAPPPAQF---PPAAVZAARQIE 1964
Qy 146 VYQAASNLVFGNNLEQTIQILDGCGGTWERTVVRALRAAYNNPERAIDYLSGIEP- 204
Db 1965 REAQOQHLYRAN-----INNGMPPG-----RDGMTPGSQTPTVGLNVPRP 2006
Qy 205 NVEAQPVARAPAAQQTN---QQAASPAQPAVALPVQSPSPASAGPNANPLNLPQGVPS 260
Db 2007 NOVSGPVMSMPPQWQQAQPIQOQPMGMPRPVMSQAQAAVAGPR-----MPN 2056
Qy 261 GGSNPGVVPVGGAGSGLDALRQL-----POFALLQLVQANPQIL----- 299
Db 2057 VQPNRSISP---SALQDLLRLTKSPSPQOQOQVNLKSNPQLMAAFIKORTAKYVANQ 2113
Qy 300 -----QPMQLGKNPQ-----ILRLIQENQAEFLRLVNESPEGPGG-NILQOLAA 346
Db 2114 PGMQPGLOSQPGMGPQPGMHQOQPSLQNLNMAQGVPRPGVPPPPQPMGSLNPGQALN 2173
Qy 347 AV-----PQTLVTPPEERAIQRLGGMGNRELVEVFFACNKBELTANTYLLDHGHEFD 401
Db 2174 IMNPGHNPMTNMPQYREMYR-----QLLQHQOQOQ 2206
Qy 402 DQQQ 405
Db 2207 QQQQ 2210

Search completed: December 17, 2003, 06:31:55
Job time : 139.842 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 06:16:40 ; Search time 14.6701 Seconds
(without alignments)
1168.082 Million cell updates/sec

Title: US-09-805-550-2

Perfect score: 2036

Sequence: 1 MKNVKTLLKGNFIEASPD.....EELTANYLLDHGHEFDQDQ 405

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/aa/PTCUS_COMB.pep.*
6: /cgn2_6/ptodata/1/aa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	2036	100.0	405	US-09-413-574-2
2	900.5	44.2	368	US-09-413-574-4
3	158.5	7.8	126	US-08-747-562-26
4	151	7.4	79	US-09-100-802-4
5	150.5	7.4	8991	US-08-714-741-32
6	141.5	6.9	2372	US-09-579-181-2
7	141.5	6.9	3118	US-09-579-181-1
8	140	6.9	2441	US-08-194-468-2
9	140	6.9	2441	US-08-961-739-2
10	140	6.9	2441	US-09-514-274-8
11	139	6.8	551	US-09-252-991A-32760
12	138	6.8	81	US-09-100-802-5
13	138	6.8	1190	US-09-107-532A-7146
14	135.5	6.7	915	US-08-480-917-2
15	135.5	6.7	915	US-09-138-736-2
16	135.5	6.7	915	US-08-988-242-2
17	135	6.6	399	US-09-252-991A-22853
18	134.5	6.6	447	US-08-450-360-4
19	134.5	6.6	486	US-08-450-360-2
20	134.5	6.6	664	US-09-328-352-6193
21	134	6.6	1088	US-09-233-857-13
22	133.5	6.6	805	US-09-103-429A-4
23	133.5	6.6	2442	US-09-514-247A-10
24	133	6.5	571	US-09-252-991A-30533
25	133	6.5	894	US-09-854-856-54
26	133	6.5	954	US-09-854-856-22
27	133	6.5	2157	US-09-854-856-52

ALIGNMENTS

RESULT 1

US-09-413-574-2
; Sequence 2, Application US/09413574
; Patent No. 6235972

; GENERAL INFORMATION:

; APPLICANT: Mahajan, Pramod B.

; APPLICANT: Tagliani, Laura

; TITLE OF INVENTION: Maize Rad23 Genes and Uses Thereof

; FILE REFERENCE: 0964

; CURRENT APPLICATION NUMBER: US/09/413,574

; EARLIER FILING DATE: 1999-10-06

; EARLIER APPLICATION NUMBER: 60/109,728

; EARLIER FILING DATE: 1998-11-23

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 2

; LENGTH: 405

; TYPE: PRT

; ORGANISM: Zea mays

US-09-413-574-2

Query Match 100.0%; Score 2036; DB 3; Length 405;
Best Local Similarity 100.0%; Pred. No. 7e-166;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MKNVKTLLKGNFIEASPDASVADVKRIIETTTQGSTYRADQQMLIYQKILKDETTLE	60
Db	1	MKNVKTLLKGNFIEASPDASVADVKRIIETTTQGSTYRADQQMLIYQKILKDETTLE	60
Qy	61	SNVAENSLVIMLSKAKASSGASTATTAKAPATLAQPAAPVAPAAASVARTPTQAPVAT	120
Db	61	SNVAENSLVIMLSKAKASSGASTATTAKAPATLAQPAAPVAPAAASVARTPTQAPVAT	120
Qy	121	AETAPPSVQPPAAATVAATDDADVYSQAASNLVFGNNLEQTQQIILDMGGTWERDVT	180
Db	121	AETAPPSVQPPAAATVAATDDADVYSQAASNLVFGNNLEQTQQIILDMGGTWERDVT	180
Qy	181	VRALRAAYNNPERRAIDLYSGIPENVEAQPVARAPAAQQTNNQQAASPAQPAVALPVQPS	240
Db	181	VRALRAAYNNPERRAIDLYSGIPENVEAQPVARAPAAQQTNNQQAASPAQPAVALPVQPS	240
Qy	241	PASAGPNANPLNLPFGVPSGSGNPGVVGAGSGALDRLQIPQFALLQLVQANPQILQ	300
Db	241	PASAGPNANPLNLPFGVPSGSGNPGVVGAGSGALDRLQIPQFALLQLVQANPQILQ	300
Qy	301	PMQLGKQNPQILRLIQENQAEFLRLVNESPEGGPGGNILGOLAAAAPQTTLVTTPERE	360
Db	301	PMQLGKQNPQILRLIQENQAEFLRLVNESPEGGPGGNILGOLAAAAPQTTLVTTPERE	360
Qy	361	AIORLECGMGNRELVLVFFACNDELTANYLLDHGHEFDQDQ 405	

Sequence 20, Appl
Sequence 50, Appl
Sequence 18, Appl
Sequence 32957, A
Sequence 3, Appl
Sequence 2, Appl
Sequence 38, Appl
Sequence 19109, A
Sequence 6, Appl
Sequence 36, Appl
Sequence 4, Appl
Sequence 34, Appl
Sequence 2, Appl
Sequence 5, Appl
Sequence 27305, A
Sequence 27660, A
Sequence 47, Appl
Sequence 64, Appl

28 133 6.5 2217 4 US-09-854-856-20
29 133 6.5 2294 4 US-09-854-856-50
30 133 6.5 2354 4 US-09-854-856-18
31 131.5 6.5 316 4 US-09-252-991A-32957
32 131.5 6.5 786 3 US-09-103-429A-3
33 131 6.4 903 2 US-08-853-310-2
34 130 6.4 922 4 US-09-854-856-38
35 130 6.4 939 4 US-09-252-991A-19109
36 130 6.4 982 4 US-09-854-856-6
37 130 6.4 2185 4 US-09-854-856-36
38 130 6.4 2245 4 US-09-854-856-4
39 130 6.4 2322 4 US-09-854-856-34
40 130 6.4 2382 4 US-09-854-856-2
41 129.5 6.4 521 4 US-09-485-286-5
42 127.5 6.3 709 4 US-09-252-991A-27305
43 126.5 6.2 531 4 US-09-252-991A-27660
44 126.5 6.2 826 4 US-09-894-998A-47
45 126.5 6.2 1911 4 US-09-854-856-64

Db 361 A I Q R L E G M G F N R E L V L E V F F A C K D E E L T A N Y L L D H G H E F D Q Q Q 405

RESULT 2

```

US-09-413-574-4
; Sequence 4, Application US/09413574
; Patent No. 6235972
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; APPLICANT: Tagliani, Laura
; TITLE OF INVENTION: Maize Rad23 Genes and Uses Thereof
; FILE REFERENCE: 0964
; CURRENT APPLICATION NUMBER: US/09/413,574
; CURRENT FILING DATE: 1999-10-06
; EARLIER APPLICATION NUMBER: 60/109,728
; EARLIER FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Zea mays
US-09-413-574-4

```

Query Match	44.2%	Score	900.5;	DB	3;	Length	368;		
Best Local Similarity	49.0%;	Pred.	No. 5.9e-69;						
Matches	197;	Conservative	53;	Mismatches	117;	Indels	35;	Gaps	7;
QY	1	MKLNVT	KLGTNFEIEBASPDASVADVKRIETTTTQGQSTYRADQQMLIYQGIKLKDTEITLLE	60					
Db	1	MKLT	TVTKLGTHTHEIRVOPNDTIMAVKNLEEIQGDSYPWGQLLIIFNGKVLKDESTLE	60					
QY	61	SNGVAENS	FIVIMLSKAKAS-SGASATATTAKAPATLIAQPAAPVAPAASVARTEPTQAPVA	119					
Db	61	ENKVNE	DGFLVMLSKGKTSGTSSOHSNTPATRQAP-----PLEAQCAP-QPPVA	114					
QY	120	TASTAPS	VOPQAAPAAATVAATDDADVISOAASNLVEGNLLEQTIOILDMGGGTWERDT	179					
Db	115	PITTSO	BEGULPACAP-----NTHDNAASNLLSGRNVDTTIINGLMEMMGGSWBKDK	164					
QY	180	VVRALRAAYNNPERAIDYLXSGIPENVEAQPVARAPAAAGQOTNOQAASPAPVALPVQP	239						
Db	165	VQRALRAAYNNPERAVEYLYSGIPVTAE----	IAPPIGGCANTTDRAFTGEA-----	213					
QY	240	SPRASAGNANPLNI	FPQGVPSGGSNPGVPFGAGSGALDALRLPQFOALLQLVOANFQIL	299					
Db	214	-GLSGI	ENTAPLDLPFGAGSNAG-----GAGGGPLDFLRNNPFQAVREMHVHNTNPQIL	266					
QY	300	QPMLOELGKONPQITRLITQENOAEBFLRVNESPEGGPGNITGLQAAAAPOTLVTPTEER	359						
Db	267	QPMVELSKONPQILRLIIEHNHDEFLQILLMEPFEGGEGGDFLDQOEDEMPHAISVTPEEQ	326						
QY	360	EAIORLEGMGFNRRELVLVEVPFACNKDBELTANYILLDRHGEPD	401						
Db	327	EATGRLSMGGFDARVTEAFACDNBELAANYILLEHAGEDB	368						

RESIST. 3

```

RESULTS 3
US-08-747-562-26
; Sequence 26, Application US/08747562
; Patent No. 6579697
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: BOLDIN, Mark
; APPLICANT: METT, Igor
; APPLICANT: VARFOLOMEJEV, Eugene
; TITLE OF INVENTION: MODULATOR OF TNF/NGF SUPERFAMILY RECEPTORS
; TITLE OF INVENTION: AND SOLUBLE OLIGOMERIC TNF/NGF SUPERFAMILY RECEPTORS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
;

```

```

; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/747,562
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05854
; FILING DATE: 11-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 109,632
; FILING DATE: 11-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 111,125
; FILING DATE: 02-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: WALLACH=15A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
US-08-747-562-26

Query Match 7.8%; Score 158.5; DB 4; Length 126;
Best Local Similarity 31.1%; Pred. No. 3e-06;
Matches 38; Conservative 25; Mismatches 50; Indels 9; Gaps 2;

QY 13 FTEASPDASVADVXRIETTGOSNYRADQQMLIYQGKILDKETTLSENGVAENSFLVI 72
DB 2 FLIRMEPDETGVKLEKIEBAEGRDAFFVAGQKLIYAGKILSDDDVPIDRYIDEKMFVVV 61
QY 73 MLSKAKA---SSSGASTATKAPATLAQPAAPVA-----PAASVARTPTQAPVATAET 123
DB 62 MVTKKAGQTSAPPEASPTAEPESSTSPAPTSCGSHPPAAPREDKSPSESTPTPEQ 121
QY 124 AP 125
DB 122 EP 123

RESULT 4
US-09-100-802-4
; Sequence 4, Application US/09100802A
; Patent No. 6294363
; GENERAL INFORMATION:
; APPLICANT: Madura, Kiran
; TITLE OF INVENTION: Methods and Compositions for the Rapid
; TITLE OF INVENTION: Purification of Proteasomes and Methods of Use of Component
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: UMDNJ97-11
; CURRENT APPLICATION NUMBER: US/09/100,802A
; CURRENT FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/050,171
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Artificial Sequence

```

```

Query Match      7.4%; Score 150.5; DB 4; Length 8991;
Best Local Similarity 21.2%; Pred. No. 0.0072;
Matches 8; Conservative 52; Mismatches 152; Indels 109; Gaps 15;

Qy 20 DASVADVKRIETQ-----GQSTYRADQOMLYQGIKILKDETTLESNGVAENSFLVI 72
      : : : : :
Db 5184 DAETAKLEKQVEDFQNSGDGEQGYLAAGSDLIKAELKFAADUK-KAVDPPE----- 5238

```

Query Match	6.9%;	Score 141.5;	DB 4;	Length 2972;
Best Local Similarity	25.5%;	Pred. No. 0.0085;		
Matches	95;	Conservative	25;	Mismatches 142; Indels 111; Gaps 16;
Qy	79	ASSGSASTATKAPAT--LAQP-AAPV-----APRASVARTTQAPVATAETAPPSVQP	130	
Db	1177	SSSLPISVPTTLAPASAPLTIPIASPLTVSAGPALLTSTVPLPAPVPAAPGPPSLQP	1236	
Qy	131	QAAPAAVTAAATDADVYSQAASLVFGNN--LSTQIQILLDMGGTWTERTVVRALRAAY	188	
Db	1237	SGASPSASALTGLGATAPSLUSSQTQPGHLLAPTSSHVPGL-----NSTVPAPACSPVL	1290	
Qy	189	NNPERAIDLYSGIPENVEAQPVARAPA-----AGQQTNOQASPAQPAVA	234	
Db	1291	VPASALASFPSPAPNPAPAQASLLAPASSASQALATPLAPMAAPQTAILAPSPAPPLAP	1349	
Qy	235	LPV-QPSASA-----GPNANP-----LMLFPQGV	258	
Db	1350	LPVLAFSPGAAPVLASSQTPVPVWAPSSGTSLASASPVPATPVLPVLPASSQTQMLPAPV	1409	
Qy	259	PS-----GGSNPGVVPGAGSGALDARQLQFOFALLQLVQANPQ	297	
Db	1410	PSPLPSPASTQTLLAPALAPLTLGGSSPQTLSLGTGN----PQGPPTQTLSLTPASSL	1465	
Qy	298	ILQP-MLQELGKNP---QILHLIQENQAEFLRLVNESPEG-----PGCGNLTGQ	343	

Db 1466 VPTPAQTLTSLAPGPGTQTLSL-----APAPPLAPASVPGPAPAHITLTAPASSASL 1520
Qy 344 LAAAVPQTLTVP 356
Db 1521 LAPASVQTLTLP 1533

RESULT 7

US-09-579-181-1
; Sequence 1, Application US/09579181
; Patent No. 6365372
; GENERAL INFORMATION:
; APPLICANT: Chiviva, John
; APPLICANT: Yaciuk, Peter
; TITLE OF INVENTION: SNF2 Related CBP Activator Protein (SRCAP)
; FILE REFERENCE: 16153-4247
; CURRENT APPLICATION NUMBER: US/09/579,181
; CURRENT FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/136,620
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3118
; TYPE: PRT
; ORGANISM: Human
US-09-579-181-1

Query Match 6.9%; Score 141.5; DB 4; Length 3118;
Best Local Similarity 25.5%; Pred. No. 0.0091;
Matches 95; Conservative 25; Mismatches 142; Indels 111; Gaps 16;
Qy 79 ASSSGSTATTAKPAT--LAQP-AAPV-----APASVARTTQAPVATAETAPPSVOP 130
Db 1323 SSSLPTISVPTLPAPASAPLTIPISAPLTVSASGPALLTSVTPPLAPVVPVAPGPPSLOP 1382
Qy 131 QAAPAAVTAATDDADVYQAASNLVFGNN--LQTTIQQILDMGGGTWERTVVRALRAAY 188
Db 1383 SGASPSALTGLATAPSLSSSQTGHPLLLAPTSSHPGL-----NSTVAPACSPVL 1436
Qy 189 NNPERAIDLYSGIPENVEAQPVARAPA-----AGQTNQOAAASPAQAVA 234
Db 1437 -VPASALASPPSAPNAPAPQAQALLAPASSAQALATPLAPMAAPQTAILAPSPAPPLAP 1495
Qy 235 LPV-QPSASA-----GPNAP-----LNLPQGV 258
Db 1496 LPVLAPSPGAAPVLAASQTVPVMAFSSTEGTSLASASVPAPTPVLAESPSTQMLPAPV 1555
Qy 259 PS-----GGSNPGVVPAGSGGALDRLQFQALLQVQNPQ 297
Db 1556 PSLPSPASTQTALAPALAPTLLGSSPSQTLSTGTN-----PQPFPTQTLSTPASSL 1611
Qy 298 ILQP-MLQELGKNP-----QILRLIQENQAEFLRLVNESPEG-----GPGNILGQ 343
Db 1612 VPTPAQTLTSLAPGPGTQTLSL-----APAPPLAPASVPGPAPAHITLTAPASSASL 1666
Qy 344 LAAAVPQTLTVP 356
Db 1667 LAPASVQTLTLP 1679

RESULT 8

US-08-194-468-2
; Sequence 2, Application US/08194468
; Patent No. 5750336
; GENERAL INFORMATION:
; APPLICANT: Montminy, Marc R.
; TITLE OF INVENTION: ASSAYS FOR THE IDENTIFICATION OF
; TITLE OF INVENTION: COMPOUNDS WHICH INHIBIT ACTIVATION OF cAMP AND MITOGEN
; TITLE OF INVENTION: RESPONSE GENES
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark

; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/194,468
; FILING DATE: 10-FEB-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9672
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)-546-4737
; TELEFAX: (619)-546-9392
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2441 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-194-468-2

Query Match 6.9%; Score 140; DB 1; Length 2441;
Best Local Similarity 21.7%; Pred. No. 0.0085;
Matches 92; Conservative 41; Mismatches 159; Indels 132; Gaps 18;
Qy 44 QMLYVQKILKDE--TTLESNGVAENSLVIMLSKAKASSSGASTATTAKAPATLAQPAAP 102
Db 1857 QHCLQAOQLMRRMATNRNVPOQSL-----PSPTAPPCTGTQQPSTPQTPOPP 1907
Qy 103 VAPAAAS-----VART--PT-----QAPVATAETAPPSVQQAAPAAVTAATDDAD 145
Db 1908 AQQPSPVNMSPAGFPNVARTQPTIVSAGKTNQVPAPPPAPQ-----PPAAVEAARQIE 1964
Qy 146 VYSQAASNLVFGNNLEQTTIQQILDMGGGTWERTVVRALRAAYNNPERAIDLYSGIPE- 204
Db 1965 REAQQQHLYRAN-----INNGMPG-----RDGMGTGSGMTVPGLNVPRP 2006
Qy 205 NYEAPQVAPAPAGQOTN-----QQAASPAQAPVALPVQSPASAGPNANPLNLPQGVPS 260
Db 2007 NQVSGFVMSMPPGQWQQAPIPQQQPMFMPRVMSMQAQAAGVGR-----MPN 2056
Qy 261 GGSNPGVVPAGSGGALDRLQ-----PQFQALLQVQNPQIL----- 299
Db 2057 VOPNRSISP-----SALQDLRLTLKSPSSPQQQQQVNLILKSNPQLMAAFIKQRTAKTVANQ 2113
Qy 300 -----QPMLELQKQNPQ-----ILRLIQENQAEFLRLVNESPEGGPGG-NILGQLAA 346
Db 2114 PGWQPGQGLSQFGMQPGMHQPGSLQNLNAMAGVPRPGVPPPPQPMAGMLNPPQOALN 2173
Qy 347 AV-----PQTLTVTPEERAIQRLSGMGFNRLVLEVPACNKDRELTYNLLDHCHEFD 401
Db 2174 INPNGHNPNMTNMPQYREVMVR-----QLLQHQOQOQ 2206
Qy 402 DQOQ 405
Db 2207 QOQO 2210

RESULT 9

US-08-961-739-2
; Sequence 2, Application US/08961739A
; Patent No. 6063583
; GENERAL INFORMATION:
; APPLICANT: Montminy, Marc R.
; TITLE OF INVENTION: Methods for Treating Diabetes Mellitus

FILE REFERENCE: SALK1650-1
CURRENT APPLICATION NUMBER: US/08/961,739A
CURRENT FILING DATE: 1997-10-31
EARLIER APPLICATION NUMBER: US 194,468
EARLIER FILING DATE: 1994-02-10
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 2441
TYPE: PRT
ORGANISM: Mus
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(2441)
OTHER INFORMATION: Xaa = Any Amino Acid
US-08-961-739-2

Query Match
Best Local Similarity 6.9%; Score 140; DB 3; Length 2441;
Matches 92; Conservative 41; Mismatches 159; Indels 132; Gaps 18;

QY 44 QMLIYQGKILKDE--TTLESNGVAENSFLVIMLSKAKASSGASTATTAKAPATLAQPAAP 102
DB 1857 QHCLQQAQLMRRMATMTRNVPOQSL-----PSPTSAPPGTPTQQPSTPTQPP 1907
QY 103 VAPAAAS-----VART--PT---QAPVATAETAPPSVQPAAPATVAATDDAD 145
DB 1908 AQOPSPVNMSPAGFPNVARTQPTIVSAGKPTNQVPAAPPAP---PPAAVEAARQIE 1964
QY 146 VYQAAASNLVFGNNLEQTIQIILDMGGTWTERTVVRALRAAYNNPRAIDLYSGIPE- 204
DB 1965 REAQOQHLYRAN-----INNMGPPG-----RDGMGTGSGQMTFVGLNVPRP 2006
QY 205 NVEAQPVARAPAAAGQTN---QQAASPAQPAVALPVQPSPASAGPNANPLNLPQGVPS 260
DB 2007 NOVSGPVNMSPGQWQAPIPQOPMGPMPRPVMSQAAQAVAGPR-----MPN 2056
QY 261 GGSNPGVVPAGSGALDALRQL-----POFQALLQVLQVANPQIL----- 299
DB 2057 VQPNRSISP---SALQDLRLTKSPSPQOQQVNLILKSNPQLMAAFIKORTAKYVANQ 2113
QY 300 -----QPMQLQELGKQNPQ-----ILRLIENQABFLRLVNESPEGPGG-NILGQLAA 346
DB 2114 PGMQOPGLOSPGQPMQPGMHQPSLQNLNAMQAGVPRGVPVPPQPMGGLNPQGOALN 2173
QY 347 AV-----PQTLTVTPEREAIQRLGEGMGNFRELVLVFFACNKKDEELTANYLLDHGHEFD 401
DB 2174 IMNPGHNPNMTNMNPQYREMYRR-----QLLQHQOQQQ 2206
QY 402 DQQQ 405
DB 2207 QQQQ 2210

RESULT 10
US-09-514-247A-8
Sequence 8, Application US/09514247A
Patent No. 6365361
GENERAL INFORMATION:
APPLICANT: TANABE SEIYAKU CO. LTD.
APPLICANT: TANIGUCHI, Tomoyasu
APPLICANT: MIZUKAMI, Junko
TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND ANTAGONIST TO PPA
FILE REFERENCE: TANIGUCHI=6
CURRENT APPLICATION NUMBER: US/09/514,247A
CURRENT FILING DATE: 2000-02-28
PRIOR APPLICATION NUMBER: PCT/JP98/03734
PRIOR FILING DATE: 1998-08-24
PRIOR APPLICATION NUMBER: JP231084/1997
PRIOR FILING DATE: 1997-08-27
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8

LENGTH: 2441
TYPE: PRT
ORGANISM: mouse
US-09-514-247A-8

Query Match
Best Local Similarity 6.9%; Score 140; DB 4; Length 2441;
Matches 92; Conservative 41; Mismatches 159; Indels 132; Gaps 18;

QY 44 QMLIYQGKILKDE--TTLESNGVAENSFLVIMLSKAKASSGASTATTAKAPATLAQPAAP 102
DB 1857 QHCLQQAQLMRRMATMTRNVPOQSL-----PSPTSAPPGTPTQQPSTPTQPP 1907
QY 103 VAPAAAS-----VART--PT---QAPVATAETAPPSVQPAAPATVAATDDAD 145
DB 1908 AQOPSPVNMSPAGFPNVARTQPTIVSAGKPTNQVPAAPPAP---PPAAVEAARQIE 1964
QY 146 VYQAAASNLVFGNNLEQTIQIILDMGGTWTERTVVRALRAAYNNPRAIDLYSGIPE- 204
DB 1965 REAQOQHLYRAN-----INNMGPPG-----RDGMGTGSGQMTFVGLNVPRP 2006
QY 205 NVEAQPVARAPAAAGQTN---QQAASPAQPAVALPVQPSPASAGPNANPLNLPQGVPS 260
DB 2007 NOVSGPVNMSPGQWQAPIPQOPMGPMPRPVMSQAAQAVAGPR-----MPN 2056
QY 261 GGSNPGVVPAGSGALDALRQL-----POFQALLQVLQVANPQIL----- 299
DB 2057 VQPNRSISP---SALQDLRLTKSPSPQOQQVNLILKSNPQLMAAFIKORTAKYVANQ 2113
QY 300 -----QPMQLQELGKQNPQ-----ILRLIENQABFLRLVNESPEGPGG-NILGQLAA 346
DB 2114 PGMQOPGLOSPGQPMQPGMHQPSLQNLNAMQAGVPRGVPVPPQPMGGLNPQGOALN 2173
QY 347 AV-----PQTLTVTPEREAIQRLGEGMGNFRELVLVFFACNKKDEELTANYLLDHGHEFD 401
DB 2174 IMNPGHNPNMTNMNPQYREMYRR-----QLLQHQOQQQ 2206
QY 402 DQQQ 405
DB 2207 QQQQ 2210

RESULT 11
US-09-252-991A-32760
Sequence 32760, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 32760
LENGTH: 551
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32760

Query Match
Best Local Similarity 6.8%; Score 139; DB 4; Length 551;
Matches 90; Conservative 58; Mismatches 129; Indels 130; Gaps 20;

QY 20 DASVADVKKRIIETQCG-----STYRADQQMLYQGLKILKDETTLENGVAENS 68
DB 2 ERNVSELIRVPDIGNGEVEIYLLVKGDKVEADQSLTLE-----SDKASME---IPSPK 54
QY 69 FLVIMLSKAKASS-----GASTATTAKAPATLAQPAAPVAPAAASVARTPTQ 115

Db 55 AGVKSIAKAGVDTLKEGGEILELEVEGEQPAEAKAERAPQAPKAE-----P 106
QY 116 APVATETAPPVQPOQAAPAAATVATDDADVTSQAASNLVFGNNLEQTTQIQLDMGGGTW 175
Db 107 AP-----APSESKAAPAAASVDIKVPDIGSAGKANVI-----EYVVRAGDVT 150
QY 176 ERDVTVRAL--RAAYNNPERAIDYLYSGIPENV-----EAQPVARAPAA 217
Db 151 EADQSILITLESKASMEISPA-----SGVSVSIVKGVDEVTGDLILUKLVEGAAPAA 205
QY 218 GOOTNOQAASPAQPAVALVPQSPASAGNPNLNFQGVPSGSGSNPGVWPCAGSGALD 277
Db 206 EE---QPAAPAAQ-AAAPAEQKPAAP-----APAKADTPAPV---GAPSRD 247
QY 278 ALRQLPQFALLQLVOANPQILQPMLOELGKNPQIL-----RLIQENQAEFLRLVNES 331
Db 248 GAK-----VHAGPAV-RMLAREFVSELSEVKASGPKGRILKEDVQVFKVKEQLQR 295
QY 332 PEGGPGGNILGQLAAAVPQTLVTPE-----EREAIQRLGGMG 369
Db 296 AKSGGAGATGG---AGIPP-----IPEVDFSKFGEVEEVANTRLMQVG 335

RESULT 12

US-09-100-802-5
; Sequence 5, Application US/09100802A
; Patent No. 6294363
; GENERAL INFORMATION:
; APPLICANT: Madura, Kiran
; TITLE OF INVENTION: Methods and Compositions for the Rapid
; TITLE OF INVENTION: Purification of Proteasomes and Methods of Use of Components
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: UMDNJ97-11
; CURRENT APPLICATION NUMBER: US/09/100,802A
; CURRENT FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/050,171
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
US-09-100-802-5

Query Match 6.8%; Score 138; DB 3; Length 81;
Best Local Similarity 36.4%; Pred. No. 8.9e-05;
Matches 28; Conservative 17; Mismatches 32; Indels 0; Gaps 0;

QY 3 LNVKTLKGNFEIEASPDASVADVKKRIETQOQSTYRADQQMLIYQGIKDKETTESN 62
Db 5 ITLTKLQQQTFKIRMEPDETVKVLKEIBAERDAPFVAGKLIYAGKILSDDDVPIRDY 64
QY 63 GVAENSLVIMLSKAKA 79
Db 65 RIDERKNFVVMVMTTKA 81

RESULT 13

US-09-107-532A-7146
; Sequence 7146, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street

CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 7146:
SEQUENCE CHARACTERISTICS:
LENGTH: 1190 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...1190
SEQUENCE DESCRIPTION: SEQ ID NO: 7146:
US-09-107-532A-7146

Query Match 6.8%; Score 138; DB 4; Length 1190;
Best Local Similarity 24.6%; Pred. No. 0.0044;
Matches 69; Conservative 29; Mismatches 112; Indels 70; Gaps 13;

QY 75 SKAKASSGASTATTAKAPATLAOPA-----APVAPAASVARTP--T 114
Db 133 SARQTAEPATGLAAAKALSRKKQFAPREDADGQTLDMFGLSSEPEQAPATPEPELS 192
QY 115 QAPVATAETAPPSVQ-----PQAAP-----AATVAATDDADVTSQAASNLVFGNNLEQTIQ 165
Db 193 QAETGASETAVPIASKPEPPQETPNIPGAAPIMAESPADRYDLGYGHM--GNGL--TW 248
QY 166 QILDMGGGTWERDVTVRALRAAYNNPERAIDYLYSGIPENV--EAQPVARAP-----AAGQ 219
Db 249 NRLEBEHDYK-----TVAHAPDRVTFTFYDADMPPEIREKTIQKVAATAEMISATQ 300
QY 220 QTNQQAASPAQPAVALP-----VQPSASAG--NNANPLNFPQGVPSGSGSNPGVV 268
Db 301 DTPVFSTPQPERVQPGNSEPEPEKVDADAANVPETGDSNTPVPSPAQKAGAP----- 357
QY 269 PGAGS-----GALDALRQLPQFQALLQLVQANPQIL 299
Db 358 TGTGSPQTEAPKTAADGLNLTPNVEEVLNKAQYDPKL 397

RESULT 14

US-08-480-917-2
; Sequence 2, Application US/08480917
; Patent No. 5820864
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: LESENECHAL, Mylene
; APPLICANT: JOLIVET, Michel
; TITLE OF INVENTION: NEW TRYPAOSOMA CRUZI ANTIGEN, AND GENE


```

; TITLE OF INVENTION: ENCODING THE LATTER; THEIR APPLICATION TO THE DETECTION OF
; TITLE OF INVENTION: CHAGAS DISEASE
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oliff & Berridge
; STREET: 700 South Washington Street, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,917
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 36400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 915 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "protein"
;
US-08-480-917-2

```

```

Query Match      6.7%; Score 135.5; DB 2; Length 915;
Best Local Similarity 23.1%; Pred. No. 0.005;
Matches 63; Conservative 43; Mismatches 126; Indels 41; Gaps 8;

Qy 71 VIMLSKAKASSSGASTATTAKAPATLAQAPAPVAPASVARTPQAPVATETAPPSVQP 130
Db 369 VVATAKPGCVSSGTDAASSHTNTTSAASAASP-----PVSAPAKAA--APPAAR 419

Qy 131 QAAPAAVTAATDDADVYSQAASNLVFGNNLEQTIQIILDMGGGTWERDVTVVALRAAYNN 190
Db 420 SAEP-----HVGSKIILANLVNQLGINVTQSVVSTGAPATTRSTAVTSTTTA--- 466

Qy 191 PERAIDYLYSGIP-----ENVEAQPVARAPAGQQTNOQAASPAQPAVALPVQPSPAS 243
Db 467 PORTSPYGHNGRPVPTAGLVANSGASAASPTAAKPTGEEKASAA-----CETSS 517

Qy 244 AGPNANPLNLFQGVPSGGSNPGVPGAGSGALDALRQLPQFQALLQLVQANPQILOPML 303
Db 518 VAINATRPALHNASLPQAPTD-GVLAATAVYQSEGEVHQ--SLERLESVITNTSRVLKLP 574

Qy 304 QELGKQNPQILRLIQENQAEFLRLVNESPEGGP 336
Db 575 DTIRRDHEQLNLGLEAQMTLQOSRPTPTQTP 607

```

```

RESULT 15
US-09-138-736-2
; Sequence 2, Application US/09138736
; Patent No. 6270767
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: LESENECHAL, Mylene
; APPLICANT: JOLIVET, Michel
; TITLE OF INVENTION: NEW TRYPANOSOMA CRUZI ANTIGEN, AND GENE
; TITLE OF INVENTION: ENCODING THE LATTER; THEIR APPLICATION TO THE DETECTION OF
; TITLE OF INVENTION: CHAGAS DISEASE
; NUMBER OF SEQUENCES: 13

```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oliff & Berridge
; STREET: 700 South Washington Street, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/138,736
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/480,917
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 36400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 915 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "protein"
;
US-09-138-736-2

```

```

Query Match      6.7%; Score 135.5; DB 3; Length 915;
Best Local Similarity 23.1%; Pred. No. 0.005;
Matches 63; Conservative 43; Mismatches 126; Indels 41; Gaps 8;

Qy 71 VIMLSKAKASSSGASTATTAKAPATLAQAPAPVAPASVARTPQAPVATETAPPSVQP 130
Db 369 VVATAKPGCVSSGTDAASSHTNTTSAASAASP-----PVSAPAKAA--APPAAR 419

Qy 131 QAAPAAVTAATDDADVYSQAASNLVFGNNLEQTIQIILDMGGGTWERDVTVVALRAAYNN 190
Db 420 SAEP-----HVGSKIILANLVNQLGINVTQSVVSTGAPATTRSTAVTSTTTA--- 466

Qy 191 PERAIDYLYSGIP-----ENVEAQPVARAPAGQQTNOQAASPAQPAVALPVQPSPAS 243
Db 467 PORTSPYGHNGRPVPTAGLVANSGASAASPTAAKPTGEEKASAA-----CETSS 517

Qy 244 AGPNANPLNLFQGVPSGGSNPGVPGAGSGALDALRQLPQFQALLQLVQANPQILOPML 303
Db 518 VAINATRPALHNASLPQAPTD-GVLAATAVYQSEGEVHQ--SLERLESVITNTSRVLKLP 574

Qy 304 QELGKQNPQILRLIQENQAEFLRLVNESPEGGP 336
Db 575 DTIRRDHEQLNLGLEAQMTLQOSRPTPTQTP 607

```

```

Search completed: December 17, 2003, 06:21:21
Job time : 15.6701 secs

```

DD 333 CONNEELLY LLENAGDFF

A:Molecule type: mRNA
A:Residues: 1-382 <STU>
A:Cross-references: EMBL:Y12013; NID:g1914682; PIDN:CAA72741.1; PID:g1914683
A:Experimental source: subspecies Queen Anne's Lace, isolate W001C
C:Genetics:
A:Gene: RAD23-1
C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

Query Match 49.9%; Score 954; DB 2; Length 382;
Best Local Similarity 52.4%; Pred. No. 1.3e-56;
Matches 205; Conservative 52; Mismatches 98; Indels 36; Gaps 9;

Qy 1 MKLTWKTLKGTFTFEIRVQPNNDTMAVKNIIEIQGKDSYPMGQQLLIIFNGKVLKDESTLE 60
Db 1 MKIYVKTLKGSQFEIQVNPDDSVADVRSIETAGAAVYPAQOQLIYQKVLKDGTTLL 60
Qy 61 ENKVNEDGFLVWMLSKGKTSSTGTSQHSNTPATRAQAPPLRAPQAP-----QPVA- 114
Db 61 ENNVAENSFVIMLSKSKSPSGSGSTTS-----TAAAPKAPQTSAPPSVPAPAVSQPPAST 116
Qy 115 -PITTSQPEGLPAQAP-----NTHNNAASNLISGRNVDTIINQLMEMGGSGWDK 163
Db 117 LPVPAPSPAPAPATAPIPSAAGSEANVYDSASLLVAGSNLEGAIQIILDMGGGTWDRD 176
Qy 164 KVQRLAAAYNNPERAVEYLYSGIPVTAETAVPI-----GGQGANTTDRAPTG--EAGLS 216
Db 177 TVIRIVRAAFNNPERAVEYLYSGIPEQAE-APPVAPSPPSGQAANPLDQPPAAQAQAPAS 235
Qy 217 GIPNAPLPLDPPQASNAG-CGAGGGGLDFLRNNPQFQAVREVMHTNPQILQPLMLVLSK 275
Db 236 AGFNANPLDPLFPQGLPDMGSSNAGAGNLDLRTNQFQALRAMVQSNFQILQPLMLBELK 295
Qy 276 QNPQILRLIENHDFLQLINEPPEGEGEDFLDQPEDEMPHAISVTPPEQEAIGRLSM 335
Db 296 QNPHMLRIQHQADFQLINPEMGGENLL-----GHGPOAISVTPPEERDAERLEAM 349
Qy 336 GFDRAVTEAFACDRNEELAANYLLEHAGE 366
Db 350 GFDRELVLVEFPACNKEELAANYLLEHME 380

RESULT 4
RAD23 protein homolog - rice
C:Species: Oryza sativa (rice)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
R:Schultz, T.F.; Quatrano, R.S.
C:Accession: T04150
A:Title: Characterization and expression of a rice RAD23 gene.
A:Reference number: 208695; MUID:97369378; PMID:9225866
A:Accession: T04150
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-392 <SCH>
A:Cross-references: EMBL:U63530; NID:g1488296; PIDN:AAB65841.1; PID:g1488297
A:Experimental source: cv. Nipponbare
C:Genetics:
A:Gene: RAD23
C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

Query Match 46.5%; Score 887.5; DB 2; Length 392;
Best Local Similarity 50.9%; Pred. No. 3.7e-52;
Matches 199; Conservative 52; Mismatches 109; Indels 31; Gaps 10;

Qy 1 MKLTWKTLKGTFTFEIRVQPNNDTMAVKNIIEIQGKDSYPMGQQLLIIFNGKVLKDESTLE 60
Db 1 MKISVTKLKGSTFQIEVDLSAQKADVVKRIETQGHYIPAEQQLIHQKVLKDDTTLD 60
Qy 61 ENKVNEDGFLVWMLSKGKTSSTGTSQHSNTPATRAQAPPLRAPQAP-----QPVAPI 118
Db 61 ENKVLNSFLVIMLRGKSGSSAPATSK-----APSNQAPPTQTPVAPASQAPVAPATT 116

A:Molecule type: DNA
A:Residues: 1-367 <STO>
A:Cross-references: GB:AE005173; NID:g7715605; PIDN:AAF68123.1; GSPDB:GN00141
C:Genetics:
A:Gene: F20B17.8
A:Map position: 1
C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

Query Match 54.6%; Score 1043.5; DB 2; Length 367;
Best Local Similarity 59.3%; Pred. No. 1.2e-62;
Matches 232; Conservative 39; Mismatches 73; Indels 47; Gaps 11;

Qy 1 MKLTWKTLKGTFTFEIRVQPNNDTMAVKNIIEIQGKDSYPMGQQLLIIFNGKVLKDESTLE 60
Db 1 MKLTWKTLKGSFHEIRVLPSPDTIMAVKNIEDSQGKNDTPCCQQLLIHNGKVLKDETSLV 60
Qy 61 ENKVNEDGFLVWMLSKGKTSSTGTSQ-----HSNTPATRAQAPPLRAPQAPQPPV 113
Db 61 ENKVTKEGFLVWMLSKSGSAGASVQCVRLLFHSLFP-----LPHRLSSIIY 111
Qy 114 APITT-----SQPEGLPAQAPNTHNNAASNLISGRNVDTIINQLMEMGGSGWDKQV 165
Db 112 NPVFTCSCTYSCSRITG-----TDYTGQAASLTVSSLSLEQVQVQIMEMGGSGWDKQV 166
Qy 166 ORALRAAYNNPERAVEYLYSGIPVTAETAVP-----ICGQGANTTDRAPTGAIGLSGIPN 220
Db 167 TRALRAAYNNPERAVDLYSGIPQTAETAVPVEAQAIGSGA-----APVAPA--SGGN 219
Qy 221 TAPLDFPQASNAGCGAGGGLDFLRNNPQFQAVREVMHTNPQILQPLMLVLSKQNPQI 280
Db 220 SSFLDLPFQETVAAGSGDLGLTLEFLRNDQVAIL--TISAFSLNCEPMLQELGKQNPQ 277
Qy 281 LRLIENHDFLQLINEPPEG--GEGDFLDQPEDEMPHAISVTPPEQEAIGRLSMGFD 338
Db 278 LRLIQENAEFLQLVNPEYEGSGEGDMFDQPEQ-EMPHAINVTPAEQAIGRLSMGFD 336
Qy 339 RARVIEAFACDRNEELAANYLLEHAGE-ED 368
Db 337 RALVIEAFACDRNEELAANYLLESGDFD 367

RESULT 3
RAD23 protein, isoform I - carrot
C:Species: Daucus carota (carrot)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
R:Sturm, A.; Leinhardt, S.
C:Accession: T14336
A:Title: Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in yeast.
A:Reference number: 217989; MUID:98345997; PMID:9681019
A:Accession: T14336
A:Status: preliminary; translated from GB/EMBL/DBJ

QY 1 MKLTVKLTGTHFEIRVQPNNDTIMAVKNIIEIQKDSYPMWQQQLIFNGKVLKDBSTLE 60
Db 1 MQVLTKTLOOQTFKIDIDPEETVKALKEIESEKGDAPFVAGQKLIYAGKILNDDTALK 60
QY 61 ENKVNEDGFLVWMLSGKT-----SGSTGTSSSOHSNTPATQAP-----PLEAPQ 106
Db 61 EYKIDENKVVVMVMTKPAVSTPAPATTQOSAPASITAVTSSITTTTVAQAPTVPALAPT 120
QY 107 QAP-----QP-PVAPITTSOPEGLPAQAP-----NTHDNA 135
Db 121 STPASITPASATASBPASAAKQEPKPAETPVATSTADTSGDSSRSNLFEDA 180
QY 136 ASNLLSGRNVDITINOLMEMGGSKDKVQRALRAAYNNPERAVEYLYSGIPVTABI-A 194
Db 181 TSALVUTQCSYENVMVTEIMSG---YEREQVIALRASFNPNDRAVEYLLMGIPGDRSQ 237
QY 195 VPIGGQGANNTDRAPTGAGLSGIPNTAPLDLPQGSNAGGAGGGLDFLRNPNPQFA 254
Db 238 VVDPPQAAT--GAPQSSAVAAAAATTT-----ATTTTSSGGHPLFLRNQPOFOQ 287
QY 255 VREVMVHTNPQILQPMVLSEKONQIILBIENHDEFLQLNPEP-----GGEGDFLD 308
Db 288 MRQIIQNPSSLPALLOQIGRENQQLQIISQHOEHFQMLNEPVGAGGQGGGGGSG 347
QY 309 QPEDEMPHA--ISVTPPEQEAIGRLSMGDFRARVTEAFIACDRNEELAANYLLEHAGE 366
Db 348 GIAEAGSGHMNYIQVTPQEKAEIRLKAIGPPEGLVITQAYFACEKNENLAANFLQQNFD 407
QY 367 ED 368
Db 408 ED 409

RESULT 8

S44443
RAD23 protein homolog2 - human
N:Alternate names: HRR23A protein
C:Species: Homo sapiens (man)
C>Date: 25-Dec-1994 #sequence_revision 01-Sep-1995 #text_change 18-Feb-2000
C:Accession: S44443; T45076
R:Masutani, C.; Sugawara, K.; Yanagisawa, J.; Sonoyama, T.; Ui, M.; Enomoto, T.; Takio, MBO J. 13, 1831-1843, 1994
A>Title: Purification and cloning of a nucleotide excision repair complex involving the A:Reference number: S44345; MUID:94222030; PMID:8168482
A:Accession: S44443
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-363 <MAS>
A:Cross-references: EMBL:D21235; NID:9498145; PIDN:BAA04767.1; PID:dl005299; PID:9498148
R:Lamerdin, J.; McCreedy, F.; Stilwogen, S.; Ramirez, M.; Carrano, A.
A:Description: Characterization by genomic sequence analysis of a gene-rich 111 kb region submitted to the EMBL Data Library, November 1996
A:Reference number: Z22906
A:Accession: T45076
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-363 <LAM>
A:Cross-references: EMBL:AD000092; PIDN:AAB51177.1
A:Experimental source: cell line 5HL2-B; fibroblast
C:Genetics:
A:Gene: RAD23A
A:Map position: 19p13.2
A:Introns: 24/3; 78/3; 139/2; 158/1; 200/3; 227/1; 271/3; 326/3
C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology
F;3-82/Domain: ubiquitin homology <UBH>

Query Match 31.6%; Score 603.5; DB 2; Length 363;
Best Local Similarity 33.4%; Pred. No. 3.7e-33;
Matches 133; Conservative 86; Mismatches 106; Indels 71; Gaps 8;
QY 3 LTVKTLKGTHTFEIRVQPNNDTIMAVKNIIEIQKDSYPMWQQQLIFNGKVLKDBSTLEEN 62
Db 5 ITLKLOOQTFKIMEPEDEIVKVLKEIEAEKGRDAFPVAGKLIYAGKILSDDDVPIRDY 64

QY 63 KVNEDGFLVWMLSGKTSSTSSQHSNTPATQAPLEAPQAPQPPVAPIT----- 117
Db 65 RIDENKVVVMVMT--KTRAGQGT-----APPEASPTAAPESSTSPFPAPTGMSPH 114
QY 118 -----TSOPEGLPAQAPNT-----HDNAASNLLSGRNVDITINOLMEMG 156
Db 115 PPAAREDKSPSESAPTTPSPSVSGSVSSGGREEDAASTLVGTSEYETMLTEIMSG 174
QY 157 GGSWMDKVKQRALRAAYNNPERAVEYLYSGIPVTABIAPVPIGGQGANNTDRAPTGAGLS 216
Db 175 ---YERERVVAALRASYNPNPRAVEYLLTGIP-----GSPEPEHGSVQBSQS 219
QY 217 GIPNTAPLDLPQGSNAGGAGGGLDFLRNPNPQFAVREVMVHTNPQILQPMVLSEK 276
Db 220 EQPATE-----AAGENPLEFLRDPQFQNMQRVITQONPALLPALLOQLQGE 285
QY 277 NPQILBIENHDEFLQLNPEP-----FEGEGDFLDQPEDEMPHAISVTPPEQEAIG 330
Db 266 NPQLLOQISRHOQFIQMLNEPPGELADISDVEGVEGAIGEAPQMYIQVTPQEKAEIR 325
QY 331 RLSEMGDFRARVTEAFIACDRNEELAANYLLEHAGEED 368
Db 326 RLKALGPPESLVITQAYFACEKNENLAANFLSQNFDE 363

RESULT 9

T40115
uv excision repair protein rad23 homolog - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Sep-2000
C:Accession: T40115; T51298
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Taylor, K.; Harris, D.
submitted to the EMBL Data Library, September 1998
A:Reference number: Z21906
A:Accession: T40115
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-368 <WOO>
A:Cross-references: EMBL:AL031788; PIDN:CAA21170.1; GSPDB:GN00067; SPDB:SPBC2D10.12
R:Zhao, Y.; Elder, R.T.
A:Experimental source: strain 972h-; cosmid c2D10
submitted to the EMBL Data Library, July 1999
A:Description: A fission yeast orthologue (rhp23) of the human nucleotide excision repair reference number: Z25362
A:Accession: T51298
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-368 <ZHA>
A:Cross-references: EMBL:AF174293; PIDN:AAD51975.1
A:Experimental source: strain SP223
C:Genetics:
A:Gene: rhp23; SPDB:SPBC2D10.12
A:Map position: 2 48/1; 328/3; 351/3
A:Introns: 23/3; 48/1; 328/3; 351/3
C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

Query Match 29.4%; Score 561.5; DB 2; Length 368;
Best Local Similarity 36.0%; Pred. No. 2.4e-30;
Matches 142; Conservative 70; Mismatches 127; Indels 55; Gaps 12;
QY 1 MKLTVKLTGTHFEIRVQPNNDTIMAVKNIIEIQKDSYPMWQQQLIFNGKVLKDBSTLE 60
Db 1 MNLTFKNLQOQKFVSDVSADTKISELK--EKIQQNYEVEVERQKLIYSGRILADDKTVG 58
QY 61 ENKVNEDGFLVWMLSGKTSSTSSQHSNTPATQAPLEAPQAPQPPVAPIT--TT 118
Db 59 EYNIKEQDFIVCMVSRPKTSTSTPKSAA----SPAPN--PPASVPBKKVEASSTVAEST 112
QY 119 SQPEGLPAQAPNTHDNAAGN-----ILSGRNVDITINOLMEMGGSKDKVKQALR 170
Db 113 STTQTVAAPSPNPDITATSEAPIDANTLAVGAQRNVAVENVMVG---YERSEVERAMR 169

[illegible]

RESULT 10
H86296
F309.1 protein - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
C/Accession: H86296
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: H86296
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-113 <STO>
A/Cross-references: GB:AE005172; NID:G94966345; PIDN:AAD34676.1; GSPDB:GN00141
C/Genetics:
A/Map position: 1

Query Match	22.2%;	Score 424;	DB 2;	Length 113;
Best Local Similarity	73.7%;	Pred. No. 8.2e-22;		
Matches	84;	Conservative 16;	Mismatches 10;	Indels 4; Gaps 3;
Qy	258	MVHTNPQILQPMLVELSKQNQPIRLRIETENHDFQLLLNEPFGGEG--DFLDPQPEDEM	315	
Dbl	1	MVASNPQILQPMLQELKQNQPLLRLIQENQAFFQLLLNEPYEGSDGDVDIFQQPDQ-EM	59	
Qy	316	PHAISVTTPQEBAIGRLSESGFDRARVIEAFLACDRNEELAANYLLEHAGE-ED	368	
Dbl	60	PHSVNVTPESEQIERLEANGGFRAIVIEAFSLCDRNEELAANYLLEHSADFED	113	

RESULT 11
T27774
hypothetical protein ZK20.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C:Accession: T27774
R;Gajadaty, S.
submitted to the EMBL Data Library, March 1996
A:Reference number: Z20417
A:Accession: T27774
A:Status: preliminary; translated from GE/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-372 <WIL>
A:cross-references: EMBL:Z69904; PIDN:CAA93780.1; GSPDB:GN00020; CBSP:ZK20.3
A:Experimental source: clone ZK20
C:Genetics:
A:Gene: CBSP:ZK20.3
A:Map position: 2

A; Introns: 49/2; 73/3; 101/3; 369/3
C; Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

Query Match 21.2%; Score 405; DB 2; Length 372;
Best Local Similarity 30.0%; Pred. No. 7.7e-20;
Matches 111; Conservative 65; Mismatches 142; Indels 52; Gaps 8

QY 1 MKLIVTKLGTHFIRVQPNDTIMAVKNIBEEIKDSYPWGQQLLIFNGKVLKDESTLE 60
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db LSVTFRTLTVNFNLENDQTIAEVKALVASEKG-DDYAPELKLIYNGKILDDSVKVG 110
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 61 ENKYNEDGFLVIMLSKGTSGTSTSSQHNTPTAQAPPLEAQQAPQPPVPAPTTSQ 120
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 111 EVGFDSKGFVVIMLSKRKVTEVAPSSTVATAAEVFPVAASP--APAADVAPEAAA- 167
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 121 PEGLPQAOPANTHDNAANLLSGRNVDTIINLMWGGSGMDKKVQRALRAAYNNPERAV 180
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 168 ---PAAEAALTDQEENVLA-----ITGMGYDRSETIAALRAAFWNPDRAV 210
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 181 EYLXSGIPVTABIAVPIGGQGANTTDRAFTGEAGLSGITPTAPLDLPFGASNAGGGAGG 240
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 211 EFLNLGLPDDA-----ADQEP-----DLGSEQINDVNDEGN 242
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 241 GPLDFLRNPQFAVRNWHVTNPOLQPMVLVELSKNQPIQLRLIEENHDEFLLQLNEPF 300
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 243 DDLNMLANMPOLAEIRALIQQNPLAAVLQOLAANVPRLVQTTQNNOQAFMIDLINGCAQ 302
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 301 GGEQDFFLDQPED-EMPHATSVTPPEOEATGRLES--MGPDFRARSVEAFACDRNEELAA 357
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 303 GAGAAAGNAPERNTPRRHVIHLSPSEAAAIERIKAVVNAPEAVVVFAYFACDKNSEAAI 362
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 358 NYLLEHAGEE 367
: : : : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 363 NFIFS NLDEE 372
: : : : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 12
S50507
excision repair protein RAD23 - yeast (Saccharomycetes cerevisiae)
N; Alternate names: protein YEL037c
C; Species: Saccharomycetes cerevisiae
C; Date: 28-Jan-1995 #sequence_revision 12-May-1995 #text_change 24-Sep-1999
C; Accession: S50507; S30845; A54599; S38547
R; Dietrich, F.S.
submitted to the EMBL Data Library, December 1994
A; Description: The sequence of S. cerevisiae cosmids 8199, 8334, and 9871.
A; Reference number: S50491
A; Accession: S50507
A; Molecule type: DNA
A; Residues: 1-398 <DIS>
A; Cross-references: EMBL:U18779; NID:g603625; PIDN:AAB65005.1; PID:g603642; MTR:
R:Mulligan, J.T.; Dietrich, F.S.; Hennessey, K.M.; Sehl, P.; Komp, C.; Wei, Y.;
submitted to the EMBL Data Library, February 1993
A; Reference number: S30812
A; Accession: S30845
A; Molecule type: DNA
A; Residues: 1-398 <MBL>
A; Cross-references: GB:L10830; NID:g603625; PIDN:AAB65005.1; PID:g
R; Watkins, J.F.; Sung, P.; Prakash, L.; Prakash, S.
Mol. Cell. Biol. 13, 7757-7765, 1993
A; Title: The Saccharomycetes cerevisiae DNA repair gene RAD23 encodes a nuclear p
A; Reference number: A54599; MUID:94067136; PMID:8246991
A; Accession: A54599
A; Molecule type: DNA
A; Residues: 1-398 <WAT>
A; Cross-references: GB:I25428; NID:g409246; PIDN:AAA16070.1; PID:g409247
J; Melnick, L.; Sherman, F.
J. Mol. Biol. 233, 372-388, 1993
A; Title: The gene clusters ARC and COR on chromosomes 5 and 10, respectively,
A; Reference number: S38543; MUID:94016558; PMID:8411151
A; Accession: S38547
A; Molecule type: DNA
A; Residues: 1-276, 'R', 278-398 <MEL>

Search completed: December 17, 2003, 06:20:49
Job time : 17.6624 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 06:16:39 ; Search time 10.4735 Seconds
(without alignments)
1652.347 Million cell updates/sec

Title: US-09-805-550-4

Perfect score: 1910

Sequence: 1 MKLTVTKLKGTHFEIRVQPN.....CDRNEELAAANYLLEHAGEED 368

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	634	33.2	416	1 R23B MOUSE	P54728 mys musculus
2	625.5	32.7	409	1 R23A HUMAN	P54727 homo sapien
3	603.5	31.6	363	1 R23B HUMAN	P54725 homo sapien
4	597.5	31.3	363	1 R23A MOUSE	P54726 mys musculus
5	561.5	29.4	368	1 RH23 SCHPO	O74803 schizosacch
6	403	21.1	398	1 RA23 YEAST	P32628 saccharomyc
7	163.5	8.6	354	1 YAUG SCHPO	Q10169 schizosacch
8	154	8.1	373	1 DSK2 YEAST	P48510 saccharomyc
9	122.5	6.4	76	1 UBIQ DROME	Q9VZ14 drosophila
10	122.5	6.4	76	1 UBIQ HUMAN	P02248 homo sapien
11	122	6.4	979	1 RFX1 HUMAN	P22670 homo sapien
12	121.5	6.4	81	1 NED8 HUMAN	Q15843 homo sapien
13	121.5	6.4	81	1 NED8 MOUSE	P29595 mys musculus
14	120.5	6.3	76	1 UBIQ EIMBO	P46574 eimeria bov
15	118.5	6.2	76	1 UBIQ LEIMA	Q05550 leishmania
16	118.5	6.2	76	1 UBIQ PHYIN	P22589 phytophthor
17	118.5	6.2	76	1 UBIQ STRPU	P23398 strongyloce
18	117.5	6.2	76	1 UBIQ LEITA	P49635 leishmania
19	116.5	6.1	76	1 UBIQ ACACA	P49634 acanthamoeb
20	116.5	6.1	76	1 UBIQ NEUCR	P31117 neurospora
21	115.5	6.0	93	1 UBIQ NPVOV	Q05120 orgyia pseu
22	114.5	6.0	76	1 UBIQ AGLNE	P42740 aglaetia pseu
23	114.5	6.0	76	1 UBIQ CAELB	P14792 caenorhabdi
24	114.5	6.0	76	1 UBIQ GEOCY	P59669 geodia cydo
25	114.5	6.0	76	1 UBIQ YEAST	P04838 saccharomyc
26	113.5	5.9	76	1 UBIQ ACECL	P42739 acetabulari
27	113.5	5.9	76	1 UBIQ CHLRE	P14624 chlamydomon
28	113.5	5.9	76	1 UBIQ COPCO	P19848 coprinus co
29	113.5	5.9	76	1 UBIQ TRYCR	P08565 trypanosoma
30	113.5	5.9	78	1 UBL1 SCHPO	Q14399 schizosacch
31	113	5.9	392	1 RME1 HUMAN	Q05925 homo sapien
32	113	5.9	989	1 FRQ NEUCR	P19970 neurospora
33	112.5	5.9	76	1 UBIQ ARATH	P59263 arabidopsis

ALIGNMENTS

RESULT 1

ID	R23B MOUSE	STANDARD;	PRT;	416 AA.
AC	P54728;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	UV excision repair protein RAD23 homolog B (MHR23B) (XP-C repair complementing repair complex 58 kDa protein) (P58).			
GN	RAD23B OR MHR23B.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BALB/c; TISSUE=Testis;			
RX	MEDLINE=96403997; PubMed=8808275;			
RA	van der Spek P.J., Vissers C.E., Hanaoka F., Smit B., Hagemeijer A., Bootsma D., Hoeijmakers J.H.J.;			
RT	"Cloning, comparative mapping, and RNA expression of the mouse homologues of the Saccharomyces cerevisiae nucleotide excision repair gene RAD23.";			
RL	Genomics 31:20-27(1996).			
CC	- - FUNCTION: INVOLVED IN DNA EXCISION REPAIR. MAY PLAY A PART IN DNA DAMAGE RECOGNITION AND/OR IN ALTERING CHROMATIN STRUCTURE TO ALLOW ACCESS BY DAMAGE-PROCESSING ENZYMES.			
CC	- - SUBUNIT: HETERODIMER OF A 125 kDa SUBUNIT (P125) AND OF A 58 kDa SUBUNIT (P58).			
CC	- - SUBCELLULAR LOCATION: Nuclear (Probable).			
CC	- - SIMILARITY: Contains 1 ubiquitin-like domain.			
CC	- - SIMILARITY: Contains 2 UBA domains.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X92411; CAA63146.1; -.			
DR	HSSP; P54725; IDV0.			
DR	MGD; MGI:105128; Rad23b.			
DR	InterPro; IPR004806; Rad23.			
DR	InterPro; IPR006636; STI1.			
DR	InterPro; IPR000449; UBA_domain.			
DR	InterPro; IPR000626; Ubiquitin.			
DR	Pfam; PF00627; UBA; 2.			
DR	Pfam; PF00240; ubiquitin; 1.			
DR	SMART; SM00165; UBA; 2.			
DR	SMART; SM00213; UBQ; 1.			
DR	TIGRFAMs; TIGR00601; rad23; 1.			
DR	PROSITE; PS0030; UBA; 2.			
DR	PROSITE; PS00053; UBIQUITIN_2; 1.			

P03993 glycine max
P20685 tetrahymena
P46379 homo sapien
P09853 herpes simp
P23324 euplates eu
P15174 trypanosoma
P08618 dictyosteli
P14283 bordetella
P77965 synechocyst
P48681 homo sapien
P45118 haemophilus
Q13495 homo sapien

34 112.5 5.9 76 1 UBIQ_SOYBN
35 112.5 5.9 76 1 UBIQ_TETPY
36 111 5.8 1132 1 BAT3_HUMAN
37 111 5.8 1144 1 RIR1_HSV23
38 110.5 5.8 76 1 UBIQ_EUPEU
39 109.5 5.7 76 1 UBIQ_TRYBB
40 108.5 5.7 76 1 UBIQ_TRYBB
41 108 5.7 910 1 PERT BORPE
42 108 5.7 1102 1 RPOB SYN3
43 107.5 5.6 1618 1 NEST_HUMAN
44 107 5.6 567 1 ODP2_HAEIN
45 106.5 5.6 701 1 CGI_HUMAN

```
KW DNA damage; DNA repair; Nuclear protein; Repeat.
FT DOMAIN 1 79 UBIQUITIN-LIKE.
FT DOMAIN 188 228 UBA 1.
FT DOMAIN 371 411 UBA 2.
FT DOMAIN 255 261 POLY-ALA.
FT DOMAIN 262 270 POLY-THR.
FT DOMAIN 336 355 POLY-GLY.
SQ SEQUENCE 416 AA; 43516 MW; 1350245AGD892205 CRC64;

Query Match 33.2%; Score 634; DB 1; Length 416;
Best Local Similarity 33.5%; Pred. No. 3.2e-33;
Matches 146; Conservative 93; Mismatches 109; Indels 88; Gaps 11;

Qy 1 MKLVTKLKGTHFIRVOPNDTIMAVKKNIEIQKDSYPMQQLLIPNGKVLKDESTLE 60
Db 1 MQVLTKLQQQTFKIDIDPEETVKALKEIESEKGDAPFVAGQKLIYAGKILSDTALK 60

Qy 61 ENKNEGDFLVMLSKGTSGTSSQHNTPA---TROAPPLEAPQ-CARPPVPAPI 116
Db 61 EYKIDENFVVMVTKRAVTTAVDATTQPSSTPSTTVSSSPAVAAQAAPATPALAPT 120

Qy 117 TT-----SOPE---GLPAQAP-----NTHDNA 135
Db 121 STPASTTASTASSEPAPAGATQEKPAEPAQTPLVLTSPADSTPGDSRSRLPEDA 180

Qy 136 ASNLISGRNVDTIINQLMEMGGGWDKQVORALRAAYNNPERAVEYLYSGIPVTAIAV 195
Db 181 TSLAVTGQSYENVMVTEISMG---YERQVIAALRASEFNNPDRAVEYLLMGIPGDRSQA 237

Qy 196 PTGGGANTTDRAPTGERAGLGIPTNAPLIDLPPQASVAG-----GGAGGGLDLFLRN 248
Db 238 VVD-----PPQAVSTTQPS-----PAVAAAAATTTATTTTSGHPLFLRN 281

Qy 249 NPQFQAVRMVHTNPQILQPMVLVELSKNPQLRLIEENHDFLQLLNEPFE----- 300
Db 282 QPQFQFQRIIQNPSLLPALLQIGRENPLQLQIQSHQEHFIQMLNEPVEAGQGQGG 341

Qy 301 -----GGEGDFLDQDEEMPHAISVTPEQEAIGRESMGFDRARVIEAFIACDRN 352
Db 342 GGGGGGGGGGGGGAEGSGHMY-IQVTPQKEAIERLKALGPPEGLVIQAYFACEKN 400

Qy 353 BELAANYLLEHAGEED 368
Db 401 ENLAANFLQQNFED 416

RESULT 2
R23B_HUMAN STANDARD; PRT; 409 AA.
AC P54727; Q8WUB0;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE UV excision repair protein RAD23 homolog B (HHR23B) (XP-C repair
DE complementing complex 58 kDa protein) (p58).
DE RAD23B.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=94222030; PubMed=8168482;
RA Masutani C., Sugawara K., Yanagisawa J., Sonoyama T., Ui M.,
RA Enomoto T., Takio K., Tanaka K., van der Spek P.J., Bootsma D.,
RA Hoeijmakers J.H.J., Hanaoka F.;
RT "Purification and cloning of a nucleotide excision repair complex
RT involving the xeroderma pigmentosum group C protein and a human
RT homologue of yeast RAD23";
RL EMBO J. 13:1831-1843(1994).
[2]
RN SEQUENCE FROM N.A., AND VARIANT VAL-249.
RP Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung M.-W.,
RA
```

```
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RA Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RA Ramsay H.;
RA Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
[4]
RN SEQUENCE FROM N.A., AND VARIANT VAL-249.
RC TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[1] FUNCTION: INVOLVED IN DNA EXCISION REPAIR. MAY PLAY A PART IN DNA
DAMAGE RECOGNITION AND/OR IN ALTERING CHROMATIN STRUCTURE TO
ALLOW ACCESS BY DAMAGE-PROCESSING ENZYMES.
[2] SUBUNIT: Heterodimer of a 125 kDa subunit (p125) and of a
58 kDa subunit (p58). Interacts with MJD.
[3] SUBCELLULAR LOCATION: Nuclear (Probable).
[4] DOMAIN: The ubiquitin-like domain mediates interaction with MJD.
[5] SIMILARITY: Contains 1 ubiquitin-like domain.
[6] SIMILARITY: Contains 2 UBA domains.
[7] This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
[8] EMBL; D21090; BAA04652.1; -
[9] EMBL; AL137852; CAD13275.1; -
[10] EMBL; AY165178; AAN47194.1; -
[11] EMBL; BC020973; AAH20973.1; -
[12] PIR; S44346; S44346.
[13] HSP; P54725; LDV0.
[14] Genew; HGNC:9813; RAD23B.
[15] MIM; 600062; -
[16] GO; GO:0005634; C:nucleus; TAS.
[17] GO; GO:0003697; F:single-stranded DNA binding activity; TAS.
[18] GO; GO:006289; P:nucleotide-excision repair; TAS.
[19] InterPro; IPR004806; Rad23.
[20] InterPro; IPR006636; STI1.
[21] InterPro; IPR000449; UBA domain.
[22] InterPro; IPR000626; Ubiquitin.
[23] Pfam; PF00627; UBA; 2.
[24] Pfam; PF00240; ubiquitin; 1.
[25] SMART; SM00727; STI1; 1.
[26] SMART; SM00165; UBA; 2.
[27] SMART; SM00213; UHQ; 1.
[28] TIGRFAMs; TIGR00601; rad23; 1.
[29] PROSITE; PS50030; UBA; 2.
[30] PROSITE; PS50053; UBIQUITIN 2; 1.
[31] DNA damage; DNA repair; Nuclear protein; Polymorphism; Repeat.
KW
```



```

DR InterPro; IPR000449; UBA domain.
DR InterPro; IPR000626; UbiQuitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiQuitin; 1.
DR SMART; SM00727; STIL; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBQ; 1.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS50053; UBIQUITIN 2; 1.
DR PROSITE; PS50053; UBIQUITIN 2; 1.
KW DNA damage; DNA repair; Nuclear protein; 3D-structure; Repeat.
FT DOMAIN 1 81 UBIQUITIN-LIKE.
FT DOMAIN 161 201 UBA 1.
FT DOMAIN 318 358 UBA 2.
FT HELIX 321 324
FT TURN 325 331
FT HELIX 334 341
FT TURN 342 345
FT HELIX 348 354
FT TURN 355
SQ SEQUENCE 363 AA; 39609 MW; C4E47B9313BB47B5 CRC64;

Query Match 31.6%; Score 603.5; DB 1; Length 363;
Best Local Similarity 33.4%; Pred. No. 2.3e-31;
Matches 133; Conservative 88; Mismatches 106; Indels 71; Gaps 8;

QY 3 LTVKTLKGTHTHEIRVQNDTIMAVKNIETIQKDSYPMGQQLIFNGKVLKDESTLEEN 62
DB 5 ITLKTQQOTFKIRMEPDETVMKLEKEAEKGRDAFPVAGQKLIYAGKILSDVDVPIRDY 64

QY 63 KVNEDGFLVMSLKGKTSSTSSQHSNTPATRAQAPLEAPQAPQAPVAPIT----- 117
DB 65 RIDENKFNVMVMTKAK-----APPEASPTAPESSTSPFPAPTSQMSHP 114

QY 118 -----TSQPEGLPAQAPNT-----HDNAASNLGSRNVDTIINQLMHWG 156
DB 115 PPAAREDKSPSEAPPTSPESVSGSVSSGSGSREEDAASTLVGSEYETMLTMSWG 174

QY 157 GSGMDKDKVQALRAAYNNPRAVEYLXGIPVTABIAVPIGGQAGTDDRAPTQAGLS 216
DB 175 -----YERERVVAALRASYNPHRAVEYLLTGIP-----GSPPEHGSVQESQVS 219

QY 217 GIPNTAPLDLFPQASNAGGAGGGLDFLRNNPQAVREVMVHTNPQILQPMVLSELSKQ 276
DB 220 EQPATE-----AAGENPLEFLRDPQFQNNRQVIQQNPALLPALLOQLQOE 265

QY 277 NQILRLIEENHDEFLLQLENEP-----PEGEGDFLDQPEDEMPHAISVTPPEQEAIG 330
DB 266 NPQLQQIISRHOEQFQMLNEPPGELADISDVEGEVGAIGEAPQNNYLIQVTPQKEALE 325

QY 331 RLESMGDFRARIETAFACDRNEELAANYLLEHAGEED 368
DB 326 RLKALGFPPESLVIQAYFACEKKNENLAANFLLSQNFDE 363

RESULT 4
R23A_MOUSE
ID R23A_MOUSE STANDARD; PRT; 363 AA.
AC P54726;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE UV excision repair protein RAD23 homolog A (MHR23A).
GN RAD23A OR MHR23A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Testis;
RX MEDLINE=96403997; PubMed=8808275;
RA van der Spek P.J., Visser C.B., Hanaoka F., Smit B.,

```

```

RA Hagemeijer A., Bootsma D., Hoeijmakers J.H.J.;
RT "Cloning, comparative mapping, and RNA expression of the mouse
RT homologues of the Saccharomyces cerevisiae nucleotide excision repair
RL gene RAD23.";
RL Genomics 31:20-27(1996).
CC -!- FUNCTION: Involved in postreplication repair of UV-damaged DNA.
CC Postreplication repair functions in gap-filling of a daughter
CC strand on replication of damaged DNA (Potential).
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: Contains 1 ubiquitin-like domain.
CC -!- SIMILARITY: Contains 2 UBA domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X92410; CAA63145.1; -.
DR HSP; P54725; LDV0.
DR MGD; MG1:105126; Rad23a.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006636; STIL.
DR InterPro; IPR000449; UBA domain.
DR InterPro; IPR000626; UbiQuitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiQuitin; 1.
DR SMART; SM00727; STIL; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBQ; 1.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS50053; UBIQUITIN 2; 1.
DR PROSITE; PS50053; UBIQUITIN 2; 1.
KW DNA damage; DNA repair; Nuclear protein; Repeat.
FT DOMAIN 1 79 UBIQUITIN-LIKE.
FT DOMAIN 161 201 UBA 1.
FT DOMAIN 318 358 UBA 2.
SQ SEQUENCE 363 AA; 39769 MW; 67EAB96EBBASC203 CRC64;

Query Match 31.3%; Score 597.5; DB 1; Length 363;
Best Local Similarity 33.7%; Pred. No. 5.5e-31;
Matches 136; Conservative 84; Mismatches 101; Indels 83; Gaps 11;

QY 3 LTVKTLKGTHTHEIRVQNDTIMAVKNIETIQKDSYPMGQQLIFNGKVLKDESTLEEN 62
DB 5 ITLKTQQOTFKIRMEPDETVMKLEKEAEKGRDAFPVAGQKLIYAGKILSDVDVPIRDY 64

QY 63 KVNEDGFLVMSLKGKTSSTSSQHSNTPATRAQAPLEAPQAPQAPVAPIT----- 116
DB 65 HIDEKFNVMVMTKAK-----APPEASPTAPESSTSPFPAPTSQMSHP 110

QY 117 -----TSQPEGLPAQAPNT-----HDNAASNLGSRNVDTIINQL 152
DB 111 MSHPPPTSRDKSPSEESTTTTSPESVSGSVSSGSGSREEDAASTLVGSEYETMLTETI 170

QY 153 MEMGGSDKDKVQALRAAYNNPRAVEYLXGIPVTABIAVPIGGQAGTDDRA--PT 210
DB 171 MSMG---YERERVVAALRASYNPHRAVEYLLTGIPGSP-----PEHG-SVQESQRAEQPA 223

QY 211 GEAGLSGIPNTAPLDLFPQASNAGGAGGGLDFLRNNPQAVREVMVHTNPQILQPMVL 270
DB 224 TEA-----AGENPLEFLRDPQFQNNRQVIQQNPALLPALLPALL 259

QY 271 VELSKNPQILRLIEENHDEFLLQLENEP-----PEGEGDFLDQPEDEMPHAISVTPPE 324
DB 260 QQLQGENPQLLQQLISRHOEQFQMLNEPPGELADISDVEGEVGAIGEAPQNNYLIQVTPQ 319

QY 325 EQEATGRLESMDFRARIETAFACDRNEELAANYLLEHAGEED 368
DB 320 EKEATERLKALGFPPESLVIQAYFACEKKNENLAANFLLSQNFDE 363

```

RESULT 5
RH23_SCHPO STANDARD; PRT; 368 AA.
ID RH23_SCHPO
AC O74803;
DR 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE UV excision repair protein rhp23 (RAD23 homolog).
GN RHP23 OR SPBC2D10.12.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
[1]
RN SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
RP STRAIN=SP223;
RC MEDLINE=21648918; PubMed=11788722;
RA Elder R.T., Song X.-Q., Chen M., Hopkins K.M., Lieberman H.B.,
RA Zhao Y.;
RT "Involvement of rhp23, a Schizosaccharomycetes pombe homolog of the
RT human HHR23A and Saccharomyces cerevisiae RAD23 nucleotide excision
RT repair genes, in cell cycle control and protein ubiquitination.";
RL Nucleic Acids Res. 30:581-591 (2002).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=972;
RC MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Waltjens J., Vansreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moesti D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardier C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsberg S.L.,
RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomycetes pombe.";
RL Nature 415:871-880 (2002).
[3]
RN CHARACTERIZATION.
RP MEDLINE=20119127; PubMed=10652237;
RA Lombaerts M., Goeloe J.I., den Dulk H., Brandsma J.A., Brouwer J.;
RT "Identification and characterization of the rhp23(+) DNA repair gene
RT in Schizosaccharomycetes pombe.";
RC Biochem. Biophys. Res. Commun. 268:210-215 (2000).
RL -1- FUNCTION: Involved in postreplication repair of UV-damaged DNA.
CC Postreplication repair functions in gap-filling of a daughter
CC strand on replication of damaged DNA.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Contains 1 ubiquitin-like domain.
CC -1- SIMILARITY: Contains 2 UBA domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

RESULT 5
RH23_SCHPO STANDARD; PRT; 368 AA.
ID RH23_SCHPO
AC O74803;
DR 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE UV excision repair protein rhp23 (RAD23 homolog).
GN RHP23 OR SPBC2D10.12.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
[1]
RN SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
RP STRAIN=SP223;
RC MEDLINE=21648918; PubMed=11788722;
RA Elder R.T., Song X.-Q., Chen M., Hopkins K.M., Lieberman H.B.,
RA Zhao Y.;
RT "Involvement of rhp23, a Schizosaccharomycetes pombe homolog of the
RT human HHR23A and Saccharomyces cerevisiae RAD23 nucleotide excision
RT repair genes, in cell cycle control and protein ubiquitination.";
RL Nucleic Acids Res. 30:581-591 (2002).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=972;
RC MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Waltjens J., Vansreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moesti D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardier C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsberg S.L.,
RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomycetes pombe.";
RL Nature 415:871-880 (2002).
[3]
RN CHARACTERIZATION.
RP MEDLINE=20119127; PubMed=10652237;
RA Lombaerts M., Goeloe J.I., den Dulk H., Brandsma J.A., Brouwer J.;
RT "Identification and characterization of the rhp23(+) DNA repair gene
RT in Schizosaccharomycetes pombe.";
RC Biochem. Biophys. Res. Commun. 268:210-215 (2000).
RL -1- FUNCTION: Involved in postreplication repair of UV-damaged DNA.
CC Postreplication repair functions in gap-filling of a daughter
CC strand on replication of damaged DNA.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Contains 1 ubiquitin-like domain.
CC -1- SIMILARITY: Contains 2 UBA domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OK NCBI_TaxID=4932;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94067136; PubMed=8246991;
 RA Watkins J.F., Sung P., Prakash L., Prakash S.;
 RT "The Saccharomycetes cerevisiae DNA repair gene RAD23 encodes a nuclear
 RT protein containing a ubiquitin-like domain required for biological
 RT function.";
 RL Mol. Cell. Biol. 13:7757-7765(1993).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=B-6441;
 RX MEDLINE=94016558; PubMed=8411151;
 RA Melnick L., Sherman F.;
 RT "The gene clusters ARC and COR on chromosomes 5 and 10, respectively,
 RT of Saccharomycetes cerevisiae share a common ancestry.";
 RL J. Mol. Biol. 233:372-388(1993).
 RN (3)
 RP SEQUENCE FROM N.A.
 RX PubMed=9169868;
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Yelton M.A., Allen E.,
 RA Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
 RA Cherry J.M., Chung E., Duncan M., Gorman E., Hartzell G.,
 RA Hunnicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
 RA Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
 RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
 RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
 RT "The nucleotide sequence of Saccharomycetes cerevisiae chromosome V.";
 RL Nature 387:78-81(1997).
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- SIMILARITY: Contains 1 ubiquitin-like domain.
 CC -!- SIMILARITY: Contains 2 UBA domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L25428; AAA16070.1; -
 DR EMBL; L22172; AAA34935.1; -
 DR EMBL; L22173; AAA34938.1; -
 DR EMBL; S65964; AAD13972.1; -
 DR EMBL; S66117; AAB28441.1; -
 DR EMBL; U18779; AAB65005.1; -
 DR PIR; S50507; S50507.
 DR HSP; P54725; IDVO.
 DR SGD; S0000763; RAD23.
 DR GO; GO:0000111; C:nucleotide excision repair factor 2 complex; IDA.
 DR GO; GO:000108; C:repairosome; IDA.
 DR GO; GO:0003684; F:damaged DNA binding activity; IDA.
 DR InterPro; IPR004806; Rad23.
 DR InterPro; IPR006636; STIL.
 DR InterPro; IPR000449; UBA domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 2.
 DR Pfam; PF00240; ubiquitin; 1.
 DR SMART; SM00727; STIL; 1.
 DR SMART; SM00165; UBA; 2.
 DR SMART; SM00213; UBQ; 1.
 DR TIGRFAMs; TIGR00601; rad23; 1.
 DR PROSITE; PS50030; UBA; 2.
 DR PROSITE; PS50053; UBQUITIN; 2; 1.
 DR PROSITE; PS50053; UBQUITIN; 2; 1.
 KW DNA damage; DNA repair; Nuclear protein; Repeat.
 FT DOMAIN 1 77 UBQUITIN-LIKE.
 FT DOMAIN 146 186 UBA 1.
 FT DOMAIN 355 395 UBA 2.
 FT DOMAIN 277 277 UBA 2.
 FT CONFLICT 277 277 A -> R (IN REF. 2).
 SQ SEQUENCE 398 AA; 42366 MW; B3F0436DAB60B833 CRC64;

Query Match 21.1%; Score 403; DB 1; Length 398;
 Best Local Similarity 31.7%; Pred. No. 1.3e-18;
 Matches 130; Conservative 63; Mismatches 157; Indels 60; Gaps 16;
 QY 1 MKLTVTKTKGTHFEIRVQPNNTIMAVK-KNIEBIOGKDSYPMQOQLIFNGKVLKDESTL 59
 DB 2 VSLTPKFKKVKPLDLPSNTILETKTKLAQISCEES---QIKLIYSGKVLQDSKTV 57
 QY 60 EENKVNEDGFLVVLMSKGTSGTSSSOHSNTPATRQAPPLEAPOQA---POPPVAPI 116
 DB 58 SEGCLKDGQDQVFWMSQKK---STKTKVTEPPAPESATTPGRENSTEASPSDASAPA 114
 QY 117 TT----SOP-EGLPQAQPNTHDNAASNLISGRNVDTIINOLMWMGGSGWKDKVQRALRA 171
 DB 115 ATAPEGQPOEQTATERTESASTGCVVGTERTETIERIMENG---YOREEVALRA 171
 QY 172 AYNNPERAVEYLYSGIPVTABIAVPIGGQGGANTTDRAPTGEAGLSGIPNTAPLDLPQGA 231
 DB 172 AFNPPRAVEYLLMGIP--ENLRQPEPQQQTAATAAEQPSATAATTAEP--AEDDLPAQA 227
 QY 232 --SNAGGA---GGPLDPLRNPP-----QFOAVREWHVNTNPOILOPMLVELSKQN 277
 DB 228 QGNNASSGALGTTGGATDAAGGPPGSGIGLTVEDLLSLRQVSGNPEALAPLENISARY 287
 QY 278 PQILRLTEENHDFLQLNPP-----PEGG-----EGDFLD-----QPBEDEM 315
 DB 288 PQUREHIMANPEVFMVSMLEAVGNMDQVMEGADDDWEGDIEVTGEEAANGLGQSGEG 347
 QY 316 PHAIVTPEQEATGRLESNGFORARVIEAFLACDRNEELAANYLL-EHA 364
 DB 348 SFQVDYTPEDDQAISSLCELGFERDLVIQVVFACDKNEAANAATILPSDHA 397
 RESULT 7
 YAUG SCHPO STANDARD; PRT; 354 AA.
 ID YAUG SCHPO
 AC Q10159;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Hypothetical protein C26A3.16 in chromosome I.
 GN SPAC26A3.16.
 OS Schizosaccharomycetes pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetaceae; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu B., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,

RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880 (2002).
 CC 1- SIMILARITY: TO YEAST DSK2.
 CC 1- SIMILARITY: Contains 1 ubiquitin-like domain.
 CC 1- SIMILARITY: Contains 1 UBA domain.
 CC 1- SIMILARITY: Contains 1 UBA domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z69240; CAA93239.1; --
 DR PIR; T38404; T38404.
 DR GeneDB SPombe; SPAC26A3.16; --
 DR InterPro; IPR006636; ST11.
 DR InterPro; IPR000449; UBA domain.
 DR InterPro; IPR006626; Ubiquitin.
 DR Pfam; PF00627; UBA; 1.
 DR Pfam; PF00240; ubiquitin; 1.
 DR SMART; SM00727; ST11; 2.
 DR SMART; SM00165; UBA; 1.
 DR SMART; SM00213; UBO; 1.
 DR PROSITE; PS50030; UBA; 1.
 DR PROSITE; PS50053; UBIQUITIN_2; 1.
 KW Hypothetical protein.
 FT DOMAIN 1 78 UBIQUITIN-LIKE.
 FT DOMAIN 309 353 UBA.
 SQ SEQUENCE 354 AA; 36819 MW; 1A99B2D97E73A831 CRC64;

Query Match 8.6%; Score 163.5; DB 1; Length 354;
 Best Local Similarity 22.7%; Pred. No. 0.0016;
 Matches 95; Conservative 62; Mismatches 134; Indels 127; Gaps 21;

QY 1 MKLTVKTLKGTIFEIRVQPNDDTMAVKNIIEIQKDSYFPGQQLLIIFNGKVLKDESTLE 60
 DB 4 ISLTIKANDQKYAVTDVSESSVLLALKEAIPVADIEK---ERQRLIYAGRVLKDSESLK 60
 QY 61 ENKVNEDGLVLMVSKGKTSGTSSQSHTNPATROAPPLEAQPOAPPPVAPITTSQ 120
 DB 61 TYKIQDGHSHILVKTLGQNPAAATVSDR-----TQVPTNIQAGQANNFLANLTSAR 115
 QY 121 PEG----LPAQA-----PNTHDNAASNLISGRNVDTIIN-----QLMEW---- 155
 DB 116 YAGFNIPMPSASMGFGNPNPVPSTTELA--NMLSNPMVQSSINEMFSNPQMLDMIINS 173
 QY 156 -----GGGSDKDKVQR-ALRAAYNPE--RAVEYLY-----SGIPVTABIAVPIGGQGA 202
 DB 174 SPHLRNAPPYVVRQMOSPEFRAMTDPDPMRQMAHQMGGAAGIDPMSLMGGGLGAGL 233
 QY 203 NTTDRAPTCEAGLGPINTAPLDLPQGSNAGGG-AGGGLDPLRNPNPQFAVREWHVT 261
 DB 234 -----GGLGAGLGFG-----GGANNATAGIAGAAPVD-----QTAAANTI--- 269
 QY 262 NPQILQPLMVELSKNPQLRLIEENHDFLOLNE-----PPEGEGD-----F 306
 DB 270 -----QN-----LLNLLGAGFGAGLGDAGLGGAGLGGAS 299
 QY 307 LQPEDENPHAIISTVPEQ--EATGRLESMGF-DRARVIEAFLACDRNEELAAANYLL 361
 DB 300 PPAPAQDTRP-----PEERYAEQLSQLNEMGFVDFERNVQALRRSGGNVQGAIESLL 351

RESULT 8

DSK2_YEAST
 ID DSK2_YEAST STANDARD; PRT; 373 AA.
 AC P48510;
 DT 01-FEB-1996 (Rel. 33, Created)

Query Match

Best Local Similarity 23.2%; Pred. No. 0.0067;
 Matches 84; Conservative 50; Mismatches 142; Indels 86; Gaps 18;

QY 1 MKLTVKTLKGTIFEIRVQPNDDTMAVKNIIEIQKDSYFPGQQLLIIFNGKVLKDESTL 59
 DB 1 MSLNIHKSQDKWEVNVAPSTVLQFKEANKANG-----IPVANORLIYSGLKLDQDTV 57

DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ubiquitin-like protein DSK2.
 GN DSK2 OR SHE4 OR YMR276W OR YMR021.02.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetaceae; Saccharomycetes;
 OC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
 NCBI TaxID:4932;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=S288c;
 RX MEDLINE=96281973; PubMed=8682868;
 RA Biggins S., Ivanovska I., Rose M.D.;
 RT "Yeast ubiquitin-like genes are involved in duplication of the
 RT microtubule organizing center.";
 RL J. Cell Biol. 133:1331-1346(1996).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RX PubMed=9169872;
 RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
 RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
 RA Jagers K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
 RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
 RT XIII.";
 RL Nature 387:90-93(1997).
 CC 1- FUNCTION: INVOLVED, WITH RAD23 IN SPINDLE POLE BODY DUPLICATION.
 CC 1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC 1- SIMILARITY: TO S.POMBE SPAC26A3.16.
 CC 1- SIMILARITY: Contains 1 ubiquitin-like domain.
 CC 1- SIMILARITY: Contains 1 UBA domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L40587; AAB07267.1; --
 DR PIR; S54583; S54583.
 DR HSP; Q15843; INDD.
 DR SGD; S0004889; DSK2.
 DR InterPro; IPR006636; ST11.
 DR InterPro; IPR000449; UBA domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 1.
 DR Pfam; PF00240; ubiquitin; 1.
 DR PRINTS; PR00348; UBIQUITIN.
 DR SMART; SM00727; ST11; 2.
 DR SMART; SM00165; UBA; 1.
 DR SMART; SM00213; UBO; 1.
 DR PROSITE; PS50030; UBA; 1.
 DR PROSITE; PS00299; UBIQUITIN_1; 1.
 DR PROSITE; PS50053; UBIQUITIN_2; 1.
 KW Nuclear protein.
 FT DOMAIN 1 76 UBIQUITIN-LIKE.
 FT DOMAIN 327 371 UBA.
 FT CONFLICT 109 109 R -> A (IN REF. 2).
 FT CONFLICT 296 296 R -> A (IN REF. 2).
 SQ SEQUENCE 373 AA; 39516 MW; 25EDF82B9DB67DF6 CRC64;

RX MEDLINE=75156547; PubMed=124018;
 RA Schlesinger D.H., Goldstein G.;
 RT "Hybrid tropinin reconstituted from vertebrate and arthropod
 subunits.";
 RL Nature 255:423-424 (1975).
 RN [7]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Mouse; STRAIN=C57BL/6;
 RX MEDLINE=90245601; PubMed=2159627;
 RA Finch J.S., Bonham K., Krieg P., Bowden G.T.;
 RT "Murine polyubiquitin mRNA sequence.";
 RL Nucleic Acids Res. 18:1907-1907 (1990).
 RN [8]
 RP SEQUENCE.
 RC SPECIES=Rat; STRAIN=Wistar; TISSUE=Duoenum;
 RX MEDLINE=94304928; PubMed=8031840;
 RA Hubbard M.J., Carne A.;
 RT "Differential feeding-related regulation of ubiquitin and
 calbindin9kDa in rat duoenum.";
 RL Biochim. Biophys. Acta 1200:191-196 (1994).
 RN [9]
 RP SEQUENCE FROM N.A. (UBA52).
 RC SPECIES=Rat; STRAIN=Sprague-Dawley;
 RX MEDLINE=96011832; PubMed=7488009;
 RA Chan Y.-L., Suzuki K., Wool I.G.;
 RT "The carboxyl extensions of two rat ubiquitin fusion proteins are
 ribosomal proteins S27a and L40.";
 RL Biochem. Biophys. Res. Commun. 215:682-690 (1995).
 RN [10]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Bovine; TISSUE=Seminale vesicle;
 RX MEDLINE=93176814; PubMed=8382528;
 RA Wempe F., Scheit K.H.;
 RT "Characterization of a full-length cDNA encoding a bovine four
 tandem-repeat ubiquitin.";
 RL Biochim. Biophys. Acta 1172:209-211 (1993).
 RN [11]
 RP SEQUENCE OF 1-74.
 RC SPECIES=Bovine;
 RX MEDLINE=75205496; PubMed=1170880;
 RA Schlesinger D.H., Goldstein G., Niall H.D.;
 RT "The complete amino acid sequence of ubiquitin, an adenylate cyclase
 stimulating polypeptide probably universal in living cells.";
 RL Biochemistry 14:2214-2218 (1975).
 RN [12]
 RP SEQUENCE OF 1-50.
 RC SPECIES=Bovine;
 RX MEDLINE=81062406; PubMed=6254502;
 RA Hamilton J.W., Rouse J.B.;
 RT "The biosynthesis of ubiquitin by parathyroid gland.";
 RL Biochem. Biophys. Res. Commun. 96:114-120 (1980).
 RN [13]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Pig;
 RX MEDLINE=88054463; PubMed=2824145;
 RA Einspanier R., Sharma H.S., Scheit K.H.;
 RT "An mRNA encoding poly-ubiquitin in porcine corpus luteum:
 identification by cDNA cloning and sequencing.";
 RL DNA 6:395-400 (1987).
 RN [14]
 RP SEQUENCE FROM N.A.
 RC SPECIES=C.porcillus; TISSUE=Spleen;
 RA Tsukagoshi N.;
 RT "Ascorbate-dependent expression of ubiquitin genes in guinea pig.";
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 RN [15]
 RP SEQUENCE FROM N.A.
 RC SPECIES=C.gribeus; STRAIN=V79; TISSUE=Lung fibroblast;
 RX MEDLINE=92223105; PubMed=1314094;
 RA Neno M., Mita K., Ichimura S.;
 RT "Evolutionarily conserved structure of the 3' non-translated region
 of a Chinese hamster polyubiquitin gene.";
 RL Biochim. Biophys. Acta 1130:247-252 (1992).

RN [16]
 RP SEQUENCE FROM N.A.
 RC SPECIES=C.longicaudatus; STRAIN=V79; TISSUE=Lung fibroblast;
 RX MEDLINE=89160248; PubMed=2537950;
 RA Fornace A.J. Jr., Alamo I. Jr., Hollander M.C., Lamoreaux E.;
 RT "Ubiquitin mRNA is a major stress-induced transcript in mammalian
 cells.";
 RL Nucleic Acids Res. 17:1215-1230 (1989).
 RN [17]
 RP SEQUENCE.
 RC SPECIES=Rabbit; TISSUE=Brain;
 RX MEDLINE=93149997; PubMed=1337207;
 RA Wajih N., Siddiqi A.R., Kaiser R., Persson B., Zaidi Z.H.,
 RA Joernvall H.;
 RT "Structural characterization of rabbit brain ubiquitin.";
 RL Protein Seq. Data Anal. 5:31-32 (1992).
 RN [18]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Chicken;
 RX MEDLINE=85213498; PubMed=2987683;
 RA Bond U., Schlesinger M.J.;
 RT "Ubiquitin is a heat shock protein in chicken embryo fibroblasts.";
 RL Mol. Cell. Biol. 5:949-956 (1985).
 RN [19]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Chicken;
 RX MEDLINE=87089821; PubMed=3025663;
 RA Bond U., Schlesinger M.J.;
 RT "The chicken ubiquitin gene contains a heat shock promoter and
 expresses an unstable mRNA in heat-shocked cells.";
 RL Mol. Cell. Biol. 6:4602-4610 (1986).
 RN [20]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Chicken; TISSUE=Spermatid;
 RX MEDLINE=89098351; PubMed=2850543;
 RA Mezquita J., Pau M., Mezquita C.;
 RT "cDNA encoding a chicken ubiquitin-fusion protein identical to the
 corresponding human protein.";
 RL Nucleic Acids Res. 16:11838-11838 (1988).
 RN [21]
 RP SEQUENCE FROM N.A.
 RC SPECIES=X.laevis;
 RX MEDLINE=85048942; PubMed=6209017;
 RA Dworkin-Rastl E., Shrutkowski A., Dworkin M.B.;
 RT "Multiple ubiquitin mRNAs during Xenopus laevis development contain
 tandem repeats of the 76 amino acid coding sequence.";
 RL Cell 39:321-325 (1984).
 RN [22]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
 RC SPECIES=Human; PubMed=3041007;
 RX MEDLINE=87311725; PubMed=3041007;
 RA Vijay-Kumar S., Bugg C.E., Cook W.J.;
 RT "Structure of ubiquitin refined at 1.8-A resolution.";
 RL J. Mol. Biol. 194:531-544 (1987).
 RN [23]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
 RC SPECIES=Human; PubMed=8166633;
 RX MEDLINE=94220022; PubMed=8166633;
 RA Ramage R., Green J., Muir T.W., Ogunjobi O.M., Love S., Shaw K.;
 RT "Synthetic, structural and biological studies of the ubiquitin system:

Query Match 6.4%; Score 122.5; DB 1; Length 76;
 Best Local Similarity 40.5%; Pred. No. 0.085;
 Matches 30; Conservative 14; Mismatches 27; Indels 3; Gaps 1;
 QY 1 MKLTVKTLKGTHFEIRVQPNDTIMAVKNTBEEIQGKDSYPWGQQLIFNGKVLKDESTLE 60
 Db 1 MQIFVKLTGTITILEVEPSDTIENVK---AKIQDKEGIPDQQLIFAGQLEDGRTLS 57
 QY 61 ENKYNEDGFLVVML 74
 Db 58 DYNIQKESTLHLVL 71

```

RESULT 11
RFX1 HUMAN STANDARD; PRT; 979 AA.
ID AC P22670;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE MHC class II regulatory factor RFX1 (RFX) (Enhancer factor C) (EF-C).
GN RFX1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RP MEDLINE=91071581; PubMed=2253877;
RA Reith W., Sanchez-Herrero C., Kober M., Silacci P., Berte C.,
RA Barras E., Mach B.;
RT "MHC class II regulatory factor RFX has a novel DNA-binding domain
RT and a functionally independent dimerization domain.";
RL Genes Dev. 4:1528-1540(1990).
[2]
IDENTITY BETWEEN RFX1 AND EF-C.
RP MEDLINE=94019311; PubMed=8413236;
RA Siegrist C.A., Durand B., Emery P., David E., Hearing P., Mach B.,
RA Reith W.;
RT "RFX1 is identical to enhancer factor C and functions as a
RT transactivator of the hepatitis B virus enhancer.";
RL Mol. Cell. Biol. 13:6375-6384(1993).
[3]
BINDING TO RPL30 PROMOTER.
RP MEDLINE=94040774; PubMed=8224874;
RA Saitany G., Perry R.P.;
RT "Transcription factor RFX1 helps control the promoter of the mouse
RT ribosomal protein-encoding gene rpl30 by binding to its alpha
RT element.";
RL Gene 132:279-283(1993).
[4]
SHOWS THAT BLS II IS NOT DUE TO RFX1.
RP MEDLINE=92375076; PubMed=1508204;
RA Sanchez-Herrero C., Reith W., Silacci P., Mach B.;
RT "The DNA-binding defect observed in major histocompatibility complex
RT class II regulatory mutants concerns only one member of a family of
RT complexes binding to the X boxes of class II promoters.";
RL Mol. Cell. Biol. 12:4076-4083(1992).
CC -1- FUNCTION: REGULATORY FACTOR ESSENTIAL FOR MHC CLASS II GENE
CC EXPRESSION. BINDS TO THE X BOXES OF MHC CLASS II GENES. ALSO BINDS
CC TO AN INVERTED REPEAT (ENH1) REQUIRED FOR HEPATITIS B VIRUS GENES
CC EXPRESSION AND TO THE MOST UPSTREAM ELEMENT (ALPHA) OF THE RPL30
CC PROMOTER.
CC -1- SUBUNIT: BINDS DNA AS AN HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE RFX FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; X58964; CAA01730.1; --
DR ENBL; A20498; CAA01506.1; --
DR PIR; A35913; A35913.
DR PDB; 1DP7; 06-MAR-00.
DR TRANSFAC; T00909; --
DR TRANSFAC; T01673; --
DR Genew; HGNC:9982; RFX1.
DR MIM; 600006; --
DR GO; GO:0003705; F:RNA polymerase II transcription factor acti. .; TAS.
DR GO; GO:0006955; P:immune response; TAS.

```

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; D23662; BAA04889.1; -.
CC PDB; 1NDD; 23-FEB-99.
CC Genew; HGNC:7732; NEDD8.
CC MIM; 603171; -.
CC GO; GO:0005634; C:nucleus; TAS.
CC GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
CC GO; GO:0006464; P:protein modification; TAS.
CC GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
CC GO; GO:0006511; P:ubiquitin-dependent protein catabolism; TAS.
CC InterPro; IPR000626; Ubiquitin.
CC Pfam; PF00240; ubiquitin; 1.
CC PRINTS; PR00348; UBIQUITIN.
CC SMART; SM00213; UBQ; 1.
CC PROSITE; PS00299; UBIQUITIN_1; 1.
CC PROSITE; PS50053; UBIQUITIN_2; 1.
KW Ub1 conjugation pathway; 3D-structure.
FT STRAND 2 7
FT TURN 8 9
FT STRAND 12 16
FT TURN 19 20
FT STRAND 22 22
FT STRAND 23 34
FT HELIX 38 40
FT STRAND 41 45
FT TURN 46 47
FT STRAND 48 49
FT TURN 52 53
FT STRAND 55 55
FT TURN 56 60
FT TURN 63 64
FT STRAND 66 71
FT TURN 74 75
SQ SEQUENCE 81 AA; 9072 MW; DC2FE102BE4725D2 CRC64;

Query Match 6.4%; Score 121.5; DB 1; Length 81;
Best Local Similarity 36.0%; Pred. No. 0.11;
Matches 27; Conservative 16; Mismatches 29; Indels 3; Gaps 1;
QY 1 MKLTVKTLGTHFEIRVQNDTIMAVKKNIEIQGKDSYPWGQQLLIFNGKVLKDESTLE 60
DB 1 MLIKVKTLTGKEIEIDIEPTDKVERKERVEE---KEGIPPOQRLLYSGKQWDEKTA 57
QY 61 ENKVNEDGFLVWMLS 75
DB 58 DYKILGSGVLHLVLA 72

RESULT 13
NED8_MOUSE
ID NED8_MOUSE STANDARD; PRT; 81 AA.
AC P29595;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Ubiquitin-like protein NEDD8.
GN NEDD8 OR NEDD-8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
CC MEDLINE=92328780; PubMed=1378265;
RA Kumar S., Tomooka Y., Noda M.;
RT "Identification of a set of genes with developmentally down-regulated expression in the mouse brain."
RL Biochem. Biophys. Res. Commun. 185:1155-1161(1992).
RN [2]

RP REVISIONS, SEQUENCE FROM N.A.
RX TISSUE=Brain;
CC MEDLINE=93371434; PubMed=8395831;
RA Kumar S., Yoshida Y., Noda M.;
RT "Cloning of a cDNA which encodes a novel ubiquitin-like protein.";
RL Biochem. Biophys. Res. Commun. 195:393-399(1993).
CC 1- FUNCTION: ACTIVATED BY AN E1-LIKE ENZYME, UBC12. THE MAJOR AND UBA3 AND THEN LINKED TO THE E2-LIKE ENZYME, UBC12. THE MAJOR TARGET PROTEIN MODIFIED BY NEDD8 IS CULLIN-4A.
CC 1- FUNCTION: MAY PLAY AN IMPORTANT ROLE DURING THE EMBRYONIC DEVELOPMENT AND DIFFERENTIATION OF THE CENTRAL NERVOUS SYSTEM. MAY PLAY AN ESSENTIAL ROLE IN EUKARYOTIC CELLULAR METABOLISM.
CC 1- TISSUE SPECIFICITY: Ubiquitous.
CC 1- DEVELOPMENTAL STAGE: DOWN-REGULATED DURING THE DEVELOPMENT OF BRAIN.
CC 1- SIMILARITY: BELONGS TO THE UBIQUITIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D10918; BAA01719.1; -.
DR FIR; JN0710; JN0710.
DR HSSP; Q15843; INDD.
DR MGD; MGI:97301; Nedd8.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR PRINTS; PR00348; UBIQUITIN.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
KW Ub1 conjugation pathway.
SQ SEQUENCE 81 AA; 8972 MW; DC339102BE4725D2 CRC64;

Query Match 6.4%; Score 121.5; DB 1; Length 81;
Best Local Similarity 36.0%; Pred. No. 0.11;
Matches 27; Conservative 16; Mismatches 29; Indels 3; Gaps 1;
QY 1 MKLTVKTLGTHFEIRVQNDTIMAVKKNIEIQGKDSYPWGQQLLIFNGKVLKDESTLE 60
DB 1 MLIKVKTLTGKEIEIDIEPTDKVERKERVEE---KEGIPPOQRLLYSGKQWDEKTA 57
QY 61 ENKVNEDGFLVWMLS 75
DB 58 DYKILGSGVLHLVLA 72
RESULT 14
UBIQ_EIMBO
ID UBIQ_EIMBO STANDARD; PRT; 76 AA.
AC P46574;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ubiquitin.
OS Eimeria bovis.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;
OC Elmeria.
OX NCBI_TaxID=5803;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93149194; PubMed=8426605;
RA Abrahamsen M.S., Clark T.G., Mascolo P., Speer C.A., White M.W.;
RT "Developmental gene expression in Eimeria bovis."
RL Mol. Biochem. Parasitol. 57:1-14(1993).
CC 1- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SELECTIVE DEGRADATION OF CELLULAR PROTEINS, THE MAINTENANCE OF CHROMATIN STRUCTURE, THE REGULATION OF GENE EXPRESSION, THE STRESS RESPONSE, AND RIBOSOME BIOGENESIS.

```
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVELS IN FULLY SPORULATED
CC OOCYSTS AND MEROZOITES. LOW LEVELS FOUND IN UNSPORULATED OOCYSTS.
CC ABSENT IN PARTIALLY SPORULATED OOCYSTS.
DR HSP; P02248; IUBI.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR PRINTS; PR00348; UBIQUITIN.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS0053; UBIQUITIN_2; 1.
KW Nuclear protein; Polyprotein.
FT SITE 48
FT BINDING 76 76
FT SEQUENCE 76 AA; 8551 MW; C59A35397EFC9B53 CRC64;
Query Match 6.3%; Score 120.5; DB 1; Length 76;
Best Local Similarity 40.5%; Pred. No. 0.11;
Matches 30; Conservative 14; Mismatches 27; Indels 3; Gaps 1;
QY 1 MKLTVKTLKGTHFEIRVQPNNDTMAVKNIIEIQGKDSYPWGQQLIFNGKVLKDESTLE 60
Db 1 MQIFVKTLGKTHFEIRVQPNNDTMAVKNIIEIQGKDSYPWGQQLIFNGKVLKDESTLE 60
QY 61 ENKVNEDGFLVVML 74
Db 58 DYNIQKESTLHLVL 71
RESULT 15
UBIQ LEIMA STANDARD; PRT; 76 AA.
AC Q05550;
DC 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ubiquitin.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]_TaxID=5664;
RP SEQUENCE FROM N.A.
RC STRAIN=MHOM/IR/83/LT252;
RX MEDLINE=93380668; PubMed=8396548;
RA Graeff G.R., Steele P.M., Peterson C.L., Bennett M.L., Langer P.J.;
RT "Sequence of a Leishmania major gene encoding an ubiquitin fusion
RT protein.";
RL Gene 131:155-156(1993).
CC -!- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SELECTIVE DEGRADATION OF
CC CELLULAR PROTEINS, THE MAINTENANCE OF CHROMATIN STRUCTURE, THE
CC REGULATION OF GENE EXPRESSION, THE STRESS RESPONSE, AND RIBOSOME
CC BIOGENESIS.
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
CC -!- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED WITH RIBOSOMAL CPEP52 AS
CC ITS C-TERMINAL EXTENSION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z14232; CAAY78598.1; -.
DR HSP; P02248; IUBI.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR PRINTS; PR00348; UBIQUITIN.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS0053; UBIQUITIN_2; 1.
```

```
KW Nuclear protein; Polyprotein.
FT SITE 48
FT BINDING 76 76
FT SEQUENCE 76 AA; 8549 MW; C52648497FFD8E47 CRC64;
Query Match 6.2%; Score 118.5; DB 1; Length 76;
Best Local Similarity 39.2%; Pred. No. 0.15;
Matches 29; Conservative 15; Mismatches 27; Indels 3; Gaps 1;
QY 1 MKLTVKTLKGTHFEIRVQPNNDTMAVKNIIEIQGKDSYPWGQQLIFNGKVLKDESTLE 60
Db 1 MQIFVKTLGKTHFEIRVQPNNDTMAVKNIIEIQGKDSYPWGQQLIFNGKVLKDESTLE 60
QY 61 ENKVNEDGFLVVML 74
Db 58 DYNIQKESTLHLVL 71
Search completed: December 17, 2003, 06:17:19
Job time : 11.4735 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 06:16:40 ; Search time 37.6093 Seconds
(without alignments)
2524.996 Million cell updates/sec

Title: US-09-805-550-4
Perfect score: 1910
Sequence: 1 MKLTVTKLKGTHPEIRVQPN.....CDRNEBLAANYLLEHAGEED 368

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTRMBL 23.1

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_podent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1169.5	61.2	365	10 Q94CE9	Q94ce9 arabidopsis
2	1169.5	61.2	371	10 Q8LA46	Q8la46 arabidopsis
3	1153.5	60.4	379	10 O03991	O03991 daucus caro
4	1043.5	54.6	367	10 Q9MA10	Q9ma10 arabidopsis
5	1019	53.4	389	10 Q9STA6	Q9ata6 lycopersico
6	978	51.2	419	10 Q9M887	Q9m887 arabidopsis
7	954	49.9	382	10 O03990	O03990 daucus caro
8	931.5	48.8	378	10 Q9FF16	Q9ff16 arabidopsis
9	930.5	48.7	378	10 Q94C35	Q94c35 arabidopsis
10	887.5	46.5	392	10 Q40742	Q40742 oryza sativ
11	661.5	34.6	246	10 Q9S9L8	Q9s9l8 arabidopsis
12	621.5	32.5	409	4 Q8WUB0	Q8wub0 homo sapien
13	599	31.4	362	11 Q8CAP3	Q8cap3 mus musculu
14	490.5	25.7	414	5 Q9XZE0	Q9xze0 drosophila
15	487.5	25.5	414	5 Q9V3W9	Q9v3w9 drosophila
16	484	25.3	341	5 Q97135	Q97135 dictyosteli

17	431	22.6	296	11 Q8BRA9	Q8bra9 mus musculu
18	424	22.2	113	10 Q9SA20	Q9sa20 arabidopsis
19	410.5	21.5	389	5 Q81JS8	Q81js8 plasmodium
20	405	21.2	372	5 Q23451	Q23451 caenorhabdi
21	405	21.2	748	6 P79370	P79370 cryctolagus
22	382	20.0	343	5 Q81MB7	Q81mb7 drosophila
23	375	19.6	290	5 Q9VCD5	Q9vcd5 drosophila
24	215	11.3	523	5 Q9NIF3	Q9nif3 dictyosteli
25	214	11.2	575	10 Q9FWF5	Q9fwf5 oryza sativ
26	204.5	10.7	142	10 Q9LF19	Q9lf19 arabidopsis
27	201	10.5	551	10 Q94C51	Q94c51 arabidopsis
28	201	10.5	551	10 Q9S1I8	Q9s1i8 arabidopsis
29	194.5	10.2	538	10 Q9S1I9	Q9s1i9 arabidopsis
30	193	10.1	65	10 Q9SCA8	Q9sca8 lycopersico
31	180.5	9.5	582	11 Q9QZM1	Q9qzm1 mus musculu
32	176	9.2	589	4 Q9H0T8	Q9h0t8 homo sapien
33	176	9.2	589	4 Q9UMX0	Q9umx0 homo sapien
34	176	9.2	589	4 Q81XS9	Q81xs9 homo sapien
35	174	9.1	595	6 Q95M59	Q95m59 bos taurus
36	171	9.0	601	4 Q9NRR5	Q9nrr5 homo sapien
37	170.5	8.9	582	11 Q9JJP9	Q9jjp9 rattus norv
38	170.5	8.9	596	11 Q9N8B8	Q9n8b8 mus musculu
39	169	8.8	314	4 Q9H8R7	Q9h8r7 homo sapien
40	167.5	8.8	555	11 Q8C7T4	Q8c7t4 mus musculu
41	167.5	8.8	582	11 Q8R317	Q8r317 mus musculu
42	167	8.7	638	11 Q9QZM0	Q9qzm0 mus musculu
43	165	8.6	589	4 Q9H3R4	Q9h3r4 homo sapien
44	163.5	8.6	502	5 Q18672	Q18672 caenorhabdi
45	160	8.4	548	10 Q8LSP9	Q8lsp9 oryza sativ

ALIGNMENTS

RESULT 1

Q94CE9 PRELIMINARY; PRT; 365 AA.
 ID Q94CE9;
 AC Q94CE9;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Putative RAD23 protein.
 GN F20B17.8.
 GE Arabidopsis thaliana (Mouse-ear cress).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,
 RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
 RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E.,
 RA Lam B., Lin J., Meyers M.C., Miranada M., Narusaka M., Nguyen M.,
 RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "Full length cDNA of gene F20B17.8 (GI:7715605).";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
 RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
 RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
 RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
 RA Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.;
 RT "Full length cDNA of gene F20B17.8 (GI:7715605).";
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY034912; AAK59419.1; -;
 DR EMBL; AY063103; AAL34277.1; -;

```

DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006636; STII.
DR InterPro; IPR000449; UBA domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00727; STII; 1.
DR SMART; SM00145; UBA; 2.
DR SMART; SM00213; UBQ; 1.
DR SMART; SM00213; UBQ; 1.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS00053; UBIQUITIN 2; 1.
SQ SEQUENCE 365 AA; 39157 MW; 4FED9EC59B467745 CRC64;

Query Match 61.2%; Score 1169.5; DB 10; Length 365;
Best Local Similarity 65.1%; Pred. No. 4.6e-75;
Matches 246; Conservative 43; Mismatches 66; Indels 23; Gaps 9;

Qy 1 MKLTIVTKLKGTHFEIRVOPNDTIMAVKKNIEIQKDSYPWGQQLLIENGKVLKDESTLE 60
Db 1 MKLTIVTKLKGTHFEIRVPSDTIMAVKKNIEDSQKDNYPGQQLLIHNGKVLKDESTLV 60

Qy 61 ENKVNEDGLVWMLSGKTSSTGTSQ--HSNTPATROAPLEAQQAPPPVAPITT 118
Db 61 ENKVTGEGFLVWMLSGKSGSAGASQVPSATTSSTKPAAP--STTQSSPVP 118

Qy 119 SQPEGLPAQAPNTHNAASLLSGRNVDTIINQLMEMGGGWDKDKVORALRAA 178
Db 119 QEQ--PAAQTDITYGQAASLLVSGSSLEQVMQVQIIMEMGGGWDKKTETVTRALRAA 175

Qy 179 AVEYLISGIPVTAETAVP-----IGGQGANTTDRAPTGEAGLSGIPNTAPLDLPQ 233
Db 176 AVDYLYSGIPQTAETAVPPEAQIAGSGA-----APVAPA--SGGPNSSPLDLFPQ 228

Qy 234 AGGAGGGGLDFLRNPNQFQAVREMVHTNPQILQMLVLSKQNPQILRLIENHDEFLQ 293
Db 229 AAGSGDLGTLFLRNNDQQLRTVHNSNPQILQMLVLSKQNPQILRLIENHDEFLQ 288

Qy 294 LINEPFG--GGDFLDQPEDEMPHAI SVTPEQEAIGRLSMGDFDRARVIEAFLACDR 351
Db 289 LVNPEYEGSDGDMFDQPEQ--EMPHAINVTPEAQEAIGRLSMGDFDRALVIEAFLACDR 347

Qy 352 NEELAANYLLEHAGE-ED 368
Db 348 NEELAANYLLENSGDFED 365

RESULT 2
Q8LA46 PRELIMINARY; PRT; 371 AA.
AC Q8LA46;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
DE DNA repair protein RAD23, putative.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX Haas B.J., Volkovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
annotation."
RL Genome Biol. 0:0-0(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY088037; AAM65583.1; -.

```

```

DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006636; STII.
DR InterPro; IPR000449; UBA domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR SMART; SM00727; STII; 1.
DR SMART; SM00145; UBA; 2.
DR SMART; SM00213; UBQ; 1.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS00053; UBIQUITIN 2; 1.
SQ SEQUENCE 371 AA; 39747 MW; 081493086EA97657 CRC64;

Query Match 61.2%; Score 1169.5; DB 10; Length 371;
Best Local Similarity 64.3%; Pred. No. 4.7e-75;
Matches 247; Conservative 43; Mismatches 65; Indels 29; Gaps 9;

Qy 1 MKLTIVTKLKGTHFEIRVOPNDTIMAVKKNIEIQKDSYPWGQQLLIENGKVLKDESTLE 60
Db 1 MKLTIVTKLKGTHFEIRVPSDTIMAVKKNIEDSQKDNYPGQQLLIHNGKVLKDESTLV 60

Qy 61 ENKVNEDGLVWMLSGKTSSTGTSQ--HSNTPATROAPLEAQQAPPPVAPITT 112
Db 61 ENKVTGEGFLVWMLSGKSGSAGASQVPSATTSSTKPAAP--STTQSSPVP 118

Qy 113 VAPITTSQPEGLPAQAPNTHNAASLLSGRNVDTIINQLMEMGGGWDKDKVORALRAA 172
Db 119 ASPIPAEQ--PAAQTDITYGQAASLLVSGSSLEQVMQVQIIMEMGGGWDKKTETVTRALRAA 175

Qy 173 YNNPERAVELYSGIPVTAETAVP-----IGGQGANTTDRAPTGEAGLSGIPNTAPLDLF 227
Db 176 YNNPERAVDYLYSGIPQTAETAVPPEAQIAGSGA-----APVAPA--SGGPNSSPLDLF 228

Qy 228 PQGASAGGGGLDFLRNPNQFQAVREMVHTNPQILQMLVLSKQNPQILRLIENHDEFLQ 287
Db 229 PQETVAAGSGDLGTLFLRNNDQQLRTVHNSNPQILQMLVLSKQNPQILRLIENHDEFLQ 288

Qy 288 HDEFLQLINEPFG--GGDFLDQPEDEMPHAI SVTPEQEAIGRLSMGDFDRARVIEA 345
Db 289 QAELQVLQVNEPYEGSDGDMFDQPEQ--EMPHAINVTPEAQEAIGRLSMGDFDRALVIEA 347

Qy 346 FLACDRNEELAANYLLEHAGE-ED 368
Db 348 FLACDRNEELAANYLLENSGDFED 371

RESULT 3
O03991 PRELIMINARY; PRT; 379 AA.
AC O03991;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
DE RAD23 protein, isoform II.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
OC Asteridae; campanulids; Apiales; Apiaceae; Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=W001C;
RX MEDLINE=98345997; PubMed=9681019;
RA Sturm A., Leinhard S.;
RT "Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in
Yeast."
RL Plant J. 13:815-821(1998).
DR EMBL; Y12014; CAA72742.1; -.
DR HSP; P54725; LDV0.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006636; STII.
DR InterPro; IPR000449; UBA domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.

```

```

DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00727; STII; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBQ; 1.
DR TIGRFAMS; TIGR00601; rad23; 1.
DR PROSITE; PS0053; UBIQUITIN 2; 1.
DR SQ SEQUENCE 379 AA; 40530 MW; B26697B439CC5929 CRC64;

Query Match
Best Local Similarity 63.0%; Score 1153.5; DB 10; Length 379;
Matches 240; Conservative 45; Mismatches 81; Indels 15; Gaps 9

Qy 1 MKLTVKTLKGTHFEIRVPNDTIMAVKKNIEBIIQKDSYPMGQQLLIFNGKYLKDESTLE 60
Db 1 MKLTVKTLKGSHFEIRAPNDTVAIAKKNIEDLQKDNVPCGQQLLIHNGKYLKDESTLA 60

Qy 61 ENKYNEDGLVYMLSKGTSSTGTGTSSSHNSNTPATROAP----PLEAQQAQPPVPVAPIT 117
Db 61 ESKI SEDGLVYMLGSKTKMSTGTGTPAAQSSAPAPTAPAVAPAPAPAPAAAPASAVIPNT 120

Qy 118 TSQPEG--LPAQAP--NTHDNAASNLGLSRNVDTIINQLMGMGGSGWDKKVQRALRAAYN 174
Db 121 T-VPEAPLSAPAPASDVTGEAASNVVAGSNLEQTTQHIMDMGGGMDTWNVSRALRAAYN 179

Qy 175 NPERAVLYLXGIPVTAIAVPIG---GGAGNTTDR--PTGEAGLS-GIPNTAFDLQFP 228
Db 180 NPERAVDLYLXGIPEMAEAAVPVSHFQDQINAGNNAISDNGVAGAA PGAPNSLFLNMFP 239

Qy 229 QGASNAGGGAGGGLDFLRNNPQFOAVTREMVHTNPQIILQPMVLVELSKONPQILRIENH 288
Db 240 QETLSGVYTGAGLSLEFLRNNPQFOTLRSWVRNPQIILQPMLELIGKONPQLLRQIQEHH 299

Qy 289 DEFLQLLNPEPEGGEGDFLDQPERDEMHPATSVTPEEQEATGRLESMSGFDRARVTEAFLA 348
Db 300 EEFQLQLNPEVASEGDMFDOPEQD-VQEQETVTAADQEAIERLEAMGFDRLGVIEAFLA 358

Qy 349 CDRNEELAAANYLLEHAGE-ED 368
Db 359 CDRNEELAVNYLLENAGDFD 379

```

RESULT 4	
Q9MA10	PRELIMINARY; PRT; 367 AA.
ID Q9MA10	
AC Q9WA10;	
DT 01-OCT-2000	(T-EMBLrel. 15, Created)
DT 01-OCT-2000	(TrEMBurel. 15, Last sequence update)
DT 01-JUN-2002	(TrEMBurel. 21, Last annotation update)
DE F20B17.8.	
OS Arabidopsis thaliana	(Mouse-ear cress).
OC Eukaryota;	Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta;	Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II;	Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;	
[1]	
RN SEQUENCE FROM N.A.	
RA Khan S., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Kim C.,	
RA Shinn P., Altafi H., Bei Q., Chin C., Chiou J., Choi E., Conn L.,	
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,	
RA Lem C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,	
RA Palz C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,	
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,	
Ecker J.R.;	
RT "Genomic sequence for Arabidopsis thaliana BAC F20B17 from chromosome	
RI I.";	
RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.	
[2]	
RN SEQUENCE FROM N.A.	
RP Ecker J.R.;	
RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.	
[3]	
RN SEQUENCE FROM N.A.	
RP Ecker J.R.;	

```

Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Alcafi H., Bai B., Chin C., Chiou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Ienz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thevari A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.,
RA Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC010793; AAF68123.1; -.
DR HSSP; P54725; 1DV0.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR000449; UBA domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBQ; 1.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS0053; UBIQUITIN 2; 1.
SQ SEQUENCE 367 AA; 39790 MW; C6D01134B5C69CE6 CRC64;

Query Match          54.6%; Score 1043.5; DB 10; Length 367;
Best Local Similarity 59.3%; Pred. No. 4.1e-66;
Matches 232; Conservative 39; Mismatches 73; Indels 47; Gaps 11

Qy 1 MKLTVTKLKGTHPEIRVOPNDTIMAVKKNIEIOGKDSYPWGQQLLIPNGKVKLDESLTE 60
Db 1 MKLTVTKLKGSHFIRVLPSDTIMAVKKNIEDSQKDYPCGQQLLIHNGKVKLDESLV 60

Qy 61 ENKYNEDGLVVLMSKGTSGTSSQ-----HSNTPATRCAPLEAPQOAPQPPV 113
Db 61 ENKYTEEGFLVVLMSKSGSGAQAQVQVRLLLFLHFFLP-----LPHLRLSIY 111

Qy 114 APIITT-----SQPEGLPAQAPNTHDNAASNLISGRNVDTIINQLMEMGGSGWDKKV 165
Db 112 NPVFTCSCTVYSCSRTTG-----TDTYQAASTLVSGSLSQMVQIINEMGGSGWDKTV 166

Qy 166 QRALRAAYNNPERAVELYSGIPVTAEIAP-----IGQGANTTDRAPTGEAGLSGIPN 220
Db 167 TRALRAAYNNPERAVDLYSGIPQTAEEVAVPVEAQIAGSGA----APVAPA--SGGPN 219

Qy 221 TAPLDLFPQASNAGGCGGLDFLRNPFQAVRENVHTNPQLQPMVLSEKONQFI 280
Db 220 SSPLDLFPQETVAAGSGDGLTLEFLRNNQDAIIL--TISAPSLNCEPMLQELGKONQL 277

Qy 281 LRLLEENHDEPLQLNPFPEG--GEGDFLDQPEDEMPHAISSVTPPEQEAIGRLSMGFD 338
Db 278 LRLIQENQAEPLQLVNSPYEGSDGEGDMFDQPEQ-EMPHAINVTPAEQEAIGRLSMGFD 336

Qy 339 RARVIEAFLACDRNEELAANYLLBHAGE-ED 368
Db 337 RALVIEAFLACDRNEELAANYLLNSGDFED 367

RESULT 5
Q9STA6 PRELIMINARY; PRT; 389 AA.
AC Q9STA6;
ID Q9STA6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE RAD23 protein.
GN RAD23.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asceridaceae; lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_taxid=4081;
[1]
RN SEQUENCE FROM N.A.

```



```

RC STRAIN=cv. West virginia 106; TISSUE=Fruit;
RA Lemaire-Chamley M., Petit J., Raymond P., Chevallier C.;
RT "Analysis of gene expression during early tomato fruit development by
RL mRNA differential display.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR HSSP; AJ243875; CAB51544.1; -.
DR HSSP; P54725; IDVO.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006636; SRI1.
DR InterPro; IPR000449; UBA_Domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00727; STI1; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBQ; 1.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS50053; UBIQUITIN 2; 1.
SQ SEQUENCE 389 AA; 41508 MW; D2BAEDF0FE70778A CRC64;

Query Match 53.4%; Score 1019; DB 10; Length 389;
Best Local Similarity 53.8%; Pred. No. 2.5e-64;
Matches 214; Conservative 47; Mismatches 97; Indels 40; Gaps 6;

Qy 1 MKLTWTKLKGTHFIRVQPNNTIMAVKNIEIQKDSYPMGQQLIFNGKVLKDESTLE 60
Db 1 MKIFVKTLKGTHFIEVKPEDSADVKNIESVQGVYPAQQLIHQGVKLDKTTILE 60

Qy 61 ENKYNEDGFLVVMLSKGTSGTSS-----SQHSNTPATRAQAPPL 102
Db 61 ENKVAENSFVIMLSKNKVS-STGTSSISAALNTAQDGTQDQRTITTFQATAALPQ 119

Qy 103 EAPQAPQPPVAPITTSQPEGLPAQPNTHNAAISLLSGHNVDTIINOLMEMGGGSDK 162
Db 120 SASESAPTPAPVPAAS-----SVTDVYDQASNLVAGSNLETTVQQLDMGGGSDR 172

Qy 163 DKVORALRAANNPERAVEYLYSGIPVTAET-----AVPIGGQGANTTDRAPT 210
Db 173 DTVRALRAANNPERAVDYLYSGIPEQTEIPPARAPAAVATAPASQAQINPAQDAS 232

Qy 211 GEAGLSGIPNTAPLDLPQGNAGGAGGGGLFLRNPNQFQAVREMVHTNPQILOPML 270
Db 233 QLAVPSSGPNANPLDLPQGLTNAGSNAGAGNLDPLRNSPQFQALRAMVQANPQILOPML 292

Qy 271 VELSKONPQIIRLIEENHDEFLLNPEPGEGBDFLDQPEDEMPHAISVTPEQEAIG 330
Db 293 QELGKONPMLRLIQEHQDFLRLINEPVE-GEQNVLQG-TAGAIPOAVTVTPEREAE 350

Qy 331 RLESGFPRARVIEAFLACDRNEELAAANYLLEHAGEED 368
Db 351 RLEMGFDRALVLEVYFACNKEELAAANYLLOHLHEFD 388

RESULT 6
QSM887 PRELIMINARY; PRT; 419 AA.
AC QSM887;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Putative RAD23 (AT3902540/F16B3_17).
GN F16B3.17.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;

```

```

RT "Arabidopsis thaliana chromosome III BAC F16B3 genomic sequence.";
RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Banh J.,
RA Bower L., Carninci P., Dale J.M., Goldsmith A.D., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yu G., Yu S., Shinozaki K.,
RA Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Shinn P., Banh J., Bower L.,
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis ORF clones.";
RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC021640; AAF32461.1; -.
RL EMBL; AV039562; AAK62617.1; -.
RL EMBL; AY113034; AAM47342.1; -.
DR HSSP; P54725; IDVO.
DR InterPro; IPR002965; P-rich_extensn.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR000449; UBA_Domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR PRINTS; PR01217; PRICHEXTENS.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS50053; UBIQUITIN 2; 1.
SQ SEQUENCE 419 AA; 44247 MW; CAAL3BC4FFEB1E25 CRC64;

Query Match 51.2%; Score 978; DB 10; Length 419;
Best Local Similarity 50.6%; Pred. No. 2.2e-61;
Matches 213; Conservative 58; Mismatches 90; Indels 60; Gaps 12;

Qy 1 MKLTWTKLKGTHFIRVQPNNTIMAVKNIEIQKDSYPMGQQLIFNGKVLKDESTLE 60
Db 1 MKIFVKTLKGTHFIEVKPEDSVVDVKNIESVQGVYPAQQLIHQGVKLDKTTIE 60

Qy 61 ENKYNEDGFLVVMLSKGT-----TSGSTGT-----SSQHS---NTPATRAQAP--- 100
Db 61 ENKVAENSFVIMMNSKPASAAAASASAGTSQAKSTPPSTQSPISQTPASVSAPVAP 120

Qy 101 -PLEAPQAPQPPVAP-----ITTSQPEGLPA-----QAP-----NTHDNAASN 138
Db 121 APTRPPTPTPTAPVAATETVTTPPEVPATISSTPTAPDSAPVGSQGVYQAASN 180

Qy 139 LLSGRNVDTIINOLMEMGGGSKDKVQALRAAYNNPERAVEYLYSGIPVTRAIAPVPI- 197
Db 181 LAAGSNLESTIQIQLDMGGGTWDRETVYALRAAFNNPERAVEYLYTGIPEQAEVP-PVA 239

Qy 198 -----GGQGANTTDRAPTGEAGLSGIPNTAPLDLPQGNAGGAGGGGLFLRNPNQF 252
Db 240 RPPASACQPNPAPPAQTOQPAAPASGPNANPLDLPQGLPNVGNPCAGTLDLRNSQOF 299

Qy 253 QAVREMVHTNPQILOPMLVELSKONPQIIRLIEENHDEFLLNPEPGEGBDFLDQ- 309
Db 300 QALRAMVQANPQILOPMLVELSKONPMLRLIQHQADFLRLINEPVEPGEGBDFLDQ- 359

Qy 310 -----PEDEMPHAISVTPEQEAIGRLSGFPRARVIEAFLACDRNEELAAANYLLEHAG 365
Db 360 AAGMPQ-----POAIQVTHEERAEIERLEANGFERALVLEVFACNKEELAAANYLLOHMH 415

Qy 366 E 366

```

```

Db          416 E 416
RESULT 7
O03990
ID O03990 PRELIMINARY; PRT; 382 AA.
AC O03990;
DT 01-JUL-1997 (TRENBLrel. 04, Created)
DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE RAD23, isoform 1.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Apiales; Apiaceae; Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W001C;
RX MEDLINE=98345997; PubMed=9681019;
RA Sturm A., Leinhardt S.;
RT "Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in
RT yeast."
RL Plant J. 13:815-821(1998).
DR EMBL; Y12013; CAA72741.1; -.
DR HSP; P54725; 1DV0.
DR InterPro; IPR002965; P-rich_extensn.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006636; STL1.
DR InterPro; IPR000449; UBA_domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR PRINTS; PR01217; PRICHTEXTENS.
DR SMART; SM00727; STL1; 1.
DR SMART; SM00165; UBA; 2.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS50053; UBQUITIN 2; 1.
SQ SEQUENCE 382 AA; 40344 MW; 27F02A2402F3DC2 CRC64;

Query Match 49.9%; Score 954; DB 10; Length 382;
Best Local Similarity 52.4%; Pred. No. 9.9e-60;
Matches 205; Conservative 52; Mismatches 98; Indels 36; Gaps 9;

QY 1 MKLVTKLKGTHFEIRVOPNDTIMAVKNIEIQKDSYPNGQQLLIINGKVLKDESTLE 60
DB 1 MKIYVTKLKGQFEIQVNPDPDSVADVKRSIETAQGAAYPAQQMLIYQGKVLKDGTTLL 60
QY 61 ENKVNEDGLVVLMSKGTSGTSGTSHNTPATROAPPLEAPQOAP-----OPVVA- 114
DB 61 ENNVAENFVIMLSKSPSGEGSTTS-----TAAAPKAPQTSAPSPVAPAVSQPPAST 116
QY 115 -PITTSQEGGLPAQAP-----NTHNAASNLGRVDTTIINQLMEMGGSGWDK 163
DB 117 LPVPAPSPAPAPATAPIPSAAVGSEANVYDAAISLLVAGSLLEGAIQIILDMGGTWD 176
QY 164 KQVRAALRAAYNNPERAVELYSGIPTVTAIEAVPI-----GGGANTTDRAPTG--EAGLS 216
DB 177 TVIRIVRAAFNNPERAVELYSGIPEQAE-APPVAPSPSGGAANPLDQPPAAAPAPAS 235
QY 217 GIPNTAPLDLPQGSASNA-GGAGGGPLDLRNNPQFOAVREWHVHTNPQIILQPMVLKSK 275
DB 236 AGPNANPLDLFPQGLPDGMSNAGNGLDFRTNQFQALRAVQSNPQIILQPMQLQGLGK 295
QY 276 QNPQILRIEENHIDEFLQLLNEPEGGEGDFLDQPEDEMPHAI SVTPPEQEAIGRLISM 335
DB 296 QNPHLMRLIQHQADFLQLINEPMEGGENLJ-----CHGQQAISVTPPEADAERLEAM 349
QY 336 GFDRARVTEAFACDRNEELAAANYLLEHAGE 366
DB 350 GFDRRLVLEVFACNKEELAANYLLDHME 380

```

```

RESULT 8
O9FF16
ID O9FF16 PRELIMINARY; PRT; 378 AA.
AC O9FF16;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE DNA repair protein RAD23 homolog (Hypothetical protein)
DE (A5938470/A5938470).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=97471969; PubMed=9330910;
RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
RA Miyajima N., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
RT features of the 1.6 Mb regions covered by twenty physically assigned
RT P1 clones."
RL DNA Res. 4:215-230(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayaashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cdna clones."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Shinn P., Banh J.,
RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayaashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis ORF clones."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation."
RL Genome Biol. 0:0-0(2002).
RN [5]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB005248; BAB09359.1; -.
DR EMBL; AY058196; AAL25609.1; -.
DR EMBL; AY081835; AAL87405.1; -.
DR EMBL; AY087564; AAM65106.1; -.
DR HSP; P54725; 1DV0.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR000449; UBA_domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.

```

Wed Dec 17 06:49:44 2003

```
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 378 AA; 40066 MW; DD9B59153EA3B6A9 CRC64;

Query Match      48.8%; Score 931.5; DB 10; Length 378;
Best Local Similarity 52.6%; Pred. No. 3.9e-58;
Matches 200; Conservative 56; Mismatches 109; Indels 15; Gaps 8;

QY 1 MKLTVTKLKTHFEIRVQNDTIMAVKNIIEIQGKDSYPMGQQLIFNGKVLKDESTLE 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MKLIFVKTLSGSNFEIEVKPADKVSIVKTAIETVKGA-YPAAKQMLIHQKVLKDETTLE 59
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 ENKVNEDGFLVYMLSGKTS---GSTGTSSSHSNTPATRQAPPLEAP-QQAPQPPVAPI 116
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 60 ENNVVENSFIVMLSKTKASPSGASTASAPASATQPTVATPQVSAPTASVPVTSGTA 119
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 117 TTSQEGPLPAQANTHDAASNLSSGRNVDTIINQLMGSGGWDKQVORALRAAYNNP 176
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 TAAAPATAASVQTDVYQGAASNLVAGTTLESTVQIILDMGGSGWDRDVTVVRALRAAFNNP 179
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 177 ERAVEYLISGIPVTAEI-----AVPIGGQGAN--TTDRAPTGEAGLSGIPNTAPLDLPQ 229
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 180 ERAVEYLISGIPPAQAEIPPVQAQAPATGEQAAANPLAQPOQAAAAPAAATGGPNANPLNLPQ 239
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 230 GASNAGGAGGGLDFLRNNPQFOAVREVMVHTNPQILQPMVLVELSKONQILRIEENHD 289
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 240 GMPAADAGAGAGNLDLRNSQQFOALRAMVQANPQILQPMQLQELKQNPQLVLIQEHQA 299
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 290 EFLQLINPEFGSGDLPDQPEDEMPHAISVTPPEQEAIGRLSEMGDFRARVIEAPLAC 349
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 300 DFLRLINEPVE-GEENVMEQLEA-AMPOAVTVPPEEREATERLEGMGDFRAMVLEVPFAC 357
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 350 DRNEELAANYLLEHAGE-ED 368
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 358 NKNEELAANYLLDHMEFED 377
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 9
Q94C35 PRELIMINARY; PRT; 378 AA.
AC Q94C35;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 40.1 kDa protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Shinn P.,
RA Tracy S.E., Ban J., Bowser L., Carninci P., Chung M.K.,
RA Goldsmith A.D., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Pham P.K., Quach H.L.,
RA Sakano H., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C.,
RA Toriumi M., Yamada K., Yu G., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY037181; AAKS9766.1; -
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006636; STI1.
DR InterPro; IPR000449; UBA domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00727; STI1; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBQ; 1.

RESULT 10
Q40742 PRELIMINARY; PRT; 392 AA.
AC Q40742;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE OsRAD23.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-Nipponbare;
RA MEDLINE=97369378; PubMed=9225866;
RA Schultz T.F., Quatrano R.S.;
RA "Characterization and expression of a rice RAD23 gene.";
RL Plant Mol. Biol. 34:557-562(1997).
DR EMBL; U63530; AAB65841.1; -
DR HSP; P54725; LDV0.
DR Gramene; Q40742; -.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006636; STI1.
DR InterPro; IPR000449; UBA domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00727; STI1; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBQ; 1.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 392 AA; 41754 MW; BD6E08574CC7C6B CRC64;
```

Query Match 46.5%; Score 887.5; DB 10; Length 392;
 Best Local Similarity 50.9%; Pred. No. 5.4e-55;
 Matches 199; Conservative 52; Mismatches 109; Indels 31; Gaps 10;

Qy 1 MKLTKVTLKGTTHPEIRVQPNNDTIMAVKKNIEEIQGKDSYPWGQQLLIIFNGKVLKDESTLE 60
 Db 1 MKISVTKLKGSTFQIEVDSQAQVADVKRIETTCQGHYPAEQQMLIHQGVKLVKDDTTLD 60

Qy 61 ENKVNEDGFLVWMLSGKTSSTSSQHSNTPATQAPLEAPQAP--OPPVAPITTT 118
 Db 61 ENKVLNEDGFLVWMLSGKTSSTSSQHSNTPATQAPLEAPQAP--OPPVAPITTT 118

Qy 119 -----SQEGLPAQAP-----NTHDNAASLLSGRNVDTIINOLMEMGGSGWMDK 163
 Db 117 VPVTVSAPPTATASPAVAVSVSEADNVGQATSLVAGSNLEATIQIILEMGGGIMDRD 176

Qy 164 KVORALRAAYNPERAVEYISGIPVTAIEAVPIGG--QCANVTTRA-----PTGEAGLSGI 218
 Db 177 IVLHALSAFNNPERAVEYISGVPQEQMDIPVPPPSIQPANPTQASQATQPAAPSILSSG 236

Qy 219 PNTAPLDFPQGASVAGGAGG--GPLDFLRNNPQAVREMYHTNPTQILQPMVLVELSKON 277
 Db 237 PNASPLDFPQALPNASTDAAGLNDALRNAQFRTLLSVQANPQILQPLLQLGLKON 296

Qy 278 PQILRLIBENHDFLQLLNEPFG--GEGDFLDQ--PEEDMPHRAISVTPEEQEAIGRLISM 335
 Db 297 PQILQLIQENQAEFLHLNPAEGDDENLLDQFPE--AMPQTAVTTEDEAILRLPEM 354

Qy 336 GFDARVTEAFIACDRNBEALANYLLEHAGE 366
 Db 355 GFDALVLDVFFACNKBQLAANYLLDHWE 385

RESULT 11

Q9S9L8
 ID Q9S9L8 PRELIMINARY; PRT; 246 AA.
 AC Q9S9L8;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE T24D18.27 protein.
 GN T24D18.27.

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Liu S.X., Yu G., Sakano H., Jhaveri A., Lee J.M., Lenz C., Pham P.,
 RA Toriumi M., Chin C., Chlou J., Choi E., Chung M., Gonzalez A.,
 RA Hwang B., Koo T., Li J., Liu A., Vayberg M., Altafi H., Brooks S.,
 RA Buehler E., Chao Q., Conn L., Conway A., Hansen N., Johnson-Hopson C.,
 RA Khan S., Kim C., Lam B., Nguyen M., Palm C., Shinn P., Tambunga G.,
 RA Davis R.W., Ecker J.R., Federpiehl N.A., Theologis A.;
 RT "The sequence of BAC T24D18 from Arabidopsis thaliana chromosome 1.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Theologis A.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC010924; AAF18513.1; -.
 DR HSSP; P02248; 1UBI.
 DR InterPro; IPR000449; UBA domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 1.
 DR Pfam; PF00240; ubiquitin; 1.
 DR SMART; SM00165; UBA; 1.
 DR SMART; SM00213; UBQ; 1.
 DR PROSITE; PS50053; UBQUITIN_2; 1.

SQ SEQUENCE 246 AA; 25966 MW; 0867D1E7D72B6FFE CRC64;
 Query Match 34.6%; Score 661.5; DB 10; Length 246;
 Best Local Similarity 56.7%; Pred. No. 3.2e-39;
 Matches 143; Conservative 29; Mismatches 63; Indels 17; Gaps 4;

Qy 1 MKLTKVTLKGTTHPEIRVQPNNDTIMAVKKNIEEIQGKDSYPWGQQLLIIFNGKVLKDESTLE 60
 Db 1 MKLTKVTLKGTTHPEIRVQPNNDTIMAVKKNIEEIQGKDSYPWGQQLLIIFNGKVLKDESTLE 60

Qy 61 ENKVNEDGFLVWMLSGKTSSTSSQHSNTPATQAPLEAPQAP--OPPVAPITTT 120
 Db 61 ENKVTBEGFLVWMLSGKTSSTSSQHSNTPATQAPLEAPQAP--OPPVAPITTT 120

Qy 121 PEGLPAPAPNTHDNAASLLSGRNVDTIINOLMEMGGSGWMDKQVORALRAAYNPERAV 180
 Db 109 VQEQPTAQSDTYGQAATLVSGSIEQVQIEMGGSGWMDKTVTRALRAAYNPERAV 168

Qy 181 EYLSGIPVTAIE--AVPIGGQANTTDRAPTGEAGLSGIPNTAPLDFPQGASVAGGAG 239
 Db 169 DYLSGIPETVTPATNLGSGVSGRELTAAPP-----SGGNSSPLDLFPQEAVIDAAGD 224

Qy 240 GGPLDFLRNNPQ 251
 Db 225 LGTLEFLRGNDQ 236

RESULT 12

Q8WUB0
 ID Q8WUB0 PRELIMINARY; PRT; 409 AA.
 AC Q8WUB0;
 DT 01-MAR-2002 (TReMBLrel. 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE RAD23 homolog B (S. cerevisiae).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Uterus;
 RA Strausberg R.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC020973; AAH20973.1; -.
 DR InterPro; IPR004806; Rad23.
 DR InterPro; IPR006636; STI1.
 DR InterPro; IPR000449; UBA domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 2.
 DR Pfam; PF00240; ubiquitin; 1.
 DR SMART; SM00727; STI1; 1.
 DR SMART; SM00165; UBA; 2.
 DR SMART; SM00213; UBQ; 1.
 DR TIGRFAMs; TIGR00601; rad23; 1.
 DR PROSITE; PS50053; UBQUITIN_2; 1.
 SQ SEQUENCE 409 AA; 43199 MW; 475FBD499DACAC69 CRC64;

Query Match 32.5%; Score 621.5; DB 4; Length 409;
 Best Local Similarity 34.6%; Pred. No. 4.5e-36;
 Matches 146; Conservative 83; Mismatches 126; Indels 67; Gaps 11;

Qy 1 MKLTKVTLKGTTHPEIRVQPNNDTIMAVKKNIEEIQGKDSYPWGQQLLIIFNGKVLKDESTLE 60
 Db 1 MQVTLKTLQOOTFKIDIDPEETVKALKKEIESEKGDAPFVAGQKLIYAGKILNDLTALK 60

Qy 61 ENKVNEDGFLVWMLSGKGT-----SGSTGTSQHSNTPATQAP--PLEAPQ 106
 Db 61 EYKIDKKNFVVMVWKPAVSTAPATQOSAPASITATVTSSTTTTVAQPTVPALAPT 120

Qy 107 QAP-----QF--PVAPITTSQEGPLPAQAP-----NTHDNA 135
 Db 121 STPASITPASATASSEPAAPASAAKQEPKPAETPVATSTATDSTSGDSSRSNLPEDA 180

QY 136 ASNLSSGRNVDITINQLMEMGGSDKQVQBALRAAANNPRAVEYLYSGIPVTAIE-A 194
 DB 181 TSALVTQSYENWVTEINSMG---VEREQVIAALRASFNPPRAVEYLLMGIPGDRBSQA 237
 QY 195 VPIGGGANTTDRAPTEAGLSGIPNTAPLDLFPQASNAGGAGGGLDFLRNNPQQA 254
 DB 238 VVDPFQAAT--GVPOSSAVAAAAATTT-----ATTTSSGGHPLEFLRNQPFQ 287
 QY 255 VREMVHTNPQILQPMVLVELSKONPQILRLIENHDEFLQLNEPFE-----GGEGDFLD 308
 DB 288 MQIIIOQNSLPPALLOQIGRENPLLOQISOHQEHIQMLNEPVEAGGCGGGGGSG 347
 QY 309 QPEEDEMPHA--ISVTPQEQAIGRLSMGFDRAVIEAFACDRNEELAANYLLEHAGE 366
 DB 348 GIAEAGSGHMYIQTVPQEKAIERLKALGFPEGLVIQAFACEKNENLAANFLQONPD 407
 QY 367 ED 368
 DB 408 ED 409

RESULT 13
 Q8CAP3 PRELIMINARY; PRT; 362 AA.
 AC Q8CAP3;
 DT 01-MAR-2003 (T-EMBLrel. 23, Created)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE RAD23a homolog.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Thymus;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium.
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT Nature 420:563-573(2002)."
 RL EMBL; AK038300; BAC29962.1; .
 SQ SEQUENCE 362 AA; 39635 MW; 804B2608ECA241A9 CRC64;

Query Match 31.4%; Score 599; DB 11; Length 362;
 Best Local Similarity 33.7%; Pred. No. 1.5e-34;
 Matches 134; Conservative 84; Mismatches 108; Indels 72; Gaps 9;

QY 3 LTVKTLKGTHTFEIRVQPNDTIMAVKNIIEIQGKDSYPMWQQLIFNGKVLKDESTLEEN 62
 DB 5 ITLKTLQOQTFKIRNEDEFTVKVKEIEAEKGRDAFPVAGQKLIYAGKILSDDVPIKEY 64
 QY 63 KVNEDGFLVNLKSGTSGTSSQHSNTPATROAPPLEAPQ-QAPQPPV----- 113
 DB 65 HIDEKNFVWVTKAKAG-----QGIPAPPEASPTAVPEPFPFVLASGMSHP 114
 QY 114 -----APITTSOPGLPAQAPNT-----HDNAANLSSGRNVDITINQLMEMG 156
 DB 115 PPTSREDKSPSESTTTTSPESIGSVSSGSSGREEDAASLTUVTSEYETMLTEINSMG 174
 QY 157 GGSWDKQVQALRAAANNPRAVEYLYSGIPVTAIEAVPIGGGANTTDRAPTGEAGLS 216
 DB 175 ---YERERVAALRASNNPRAVEYLLTGIP-----GSPEPEHSGVQESAP----- 219
 QY 217 GIPNTAPLDLFPQASNAGGAGGGLDFLRNNPQQAVERMVHTNPQILQPMVLVELSKQ 276
 DB 220 -----EQPATE---AGENPLEFLRDQPFQNMQRQVIOQNPAALLPALLQQLGQE 264
 QY 277 NPQILRLIENHDEFLQLNEP-----FEGEGDFLDQPEDEMPHAISVTPQEQAIG 330
 DB 265 NPQLLQIISRHQEQFIQMLNEPPEGLADISDVEGEVGAIGEAPQMNYIQTVPQEKAE 324

QY 331 RLESMGFDRAVIEAFACDRNEELAANYLLEHAGEBD 368
 DB 325 RLKALGPESLVQIYAFACEKNENLAANFLLSQNFDE 362

RESULT 14
 Q9XZE0 PRELIMINARY; PRT; 414 AA.
 ID Q9XZE0;
 AC Q9XZE0;
 DT 01-NOV-1999 (T-EMBLrel. 12, Created)
 DT 01-NOV-1999 (T-EMBLrel. 12, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE DR23.
 GN RAD23 OR DHR23 OR CG1836.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nabirochikina E.N., Grischuk A.L., Soldatov A.V.;
 RT "Cloning and characterization of the Drosophila melanogaster homologue
 RT of the Saccharomyces cerevisiae gene RAD23.";
 RL Genetika 35:0-0(1999).
 DR EMBL; AP136606; AAD33695.1; .
 DR HSSP; P54725; IDV0.
 DR FlyBase; FBgn0026777; Rad23.
 DR InterPro; IPR004806; Rad23.
 DR InterPro; IPR006636; STI1.
 DR InterPro; IPR000449; UBA domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 2.
 DR Pfam; PF00240; ubiquitin; 1.
 DR SMART; SM00727; STI1; 1.
 DR SMART; SM00165; UBA; 2.
 DR SMART; SM00213; UBQ; 1; rad23; 1.
 DR TIGRFBME; TIGR00601; rad23; 1.
 DR PROSITE; PS50053; UBIQUITIN 2; 1.
 SQ SEQUENCE 414 AA; 45794 MW; 4C2FF94C116F7AB CRC64;

Query Match 25.7%; Score 490.5; DB 5; Length 414;
 Best Local Similarity 29.0%; Pred. No. 9.2e-27;
 Matches 129; Conservative 77; Mismatches 130; Indels 109; Gaps 11;

QY 1 MKLTVKTLKGTHTFEIRVQPNDTIMAVKNIIEIQGKDSYPMWQQLIFNGKVLKDESTLE 60
 DB 1 MIITIKLQOQTFITIEFAPEKTVLELKKKIFEEERGEPE-YVAEKQKLIYAGVILTDORTVG 59
 QY 61 ENKVNEDGFLVNLKSGTSGTSSQHSN-----TPATROAPPLEAPQAPQP----- 111
 DB 60 SYNVDKKFIVMLTRDSSSNRNQLSVKESNKLSTDDSKQSMPCERANHTNSPSTNT 119
 QY 112 -----PVAPITTSQPEGLPAQAPNTHDNAANLSSGRNVDITINQLMEMGGSDKQK 164
 DB 120 EDSVLSRETRPUSDDSLIGELAQ-SLQSRASNLMDGEYNTVLSWEMG---YPREQ 175
 QY 165 VQALRAAANNPRAVEYLYSGIPV-----TAEIAPVPIGGGANTTDRAPTGE 212
 DB 176 VERAMPAASNNPRAVEYLYINGIPAEEGTFYRNELNSTNPISLPISGFPASATSAERSTE 235
 QY 213 AGLSGIPNTAPLDLFPQASNAGGAGGGLDFLRNNPQQAVERMVHTNPQILQPMVLVE 272
 DB 236 -----SNSDPFEFLRSQPFQFLQMRSLIYQNPHLLHVLQ 270
 QY 273 LSKONPQILRLIENHDEFLQLNEP-----EGG----- 302
 DB 271 IGTNPALLQIISRHQEQFIQMLNEPPEGLADISDVEGEVGAIGEAPQMNYIQTVPQEKAE 330
 QY 303 -----EGDFLDQPEDEMPHAIS---VTPQEQAIGRLSMGFDRAV 342
 DB 331 VATSAQRSAGTSNAHSGSAAADN-EDLEQPLGVSTIRLNQDKDAIERLKALGFPEALV 389

```

Qy 343 IEAFLACDRNEELAANYLLEHAGEE 367
Db 390 LQAYFACEKNEEQANFLSSSFED 414

RESULT 15
Q9V3W9 PRELIMINARY; PRT; 414 AA.
AC Q9V3W9;
DT 01-MAY-2000 (TreeBLrel. 13, Created)
DT 01-MAY-2000 (TreeBLrel. 13, Last sequence update)
DE 01-MAR-2003 (TreeBLrel. 23, Last annotation update)
DE RAD23 protein.
GN RAD23 OR CG1836.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazew R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Fowler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Brodsky M.H., Rubin G.M., Tsang G.;
RT "Full Length Drosophila melanogaster cDNA sequence.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003844; AAF59352.1;
DR EMBL; AF132147; AAD33594.1;
DR HSSP; P54725; IDV0.
DR FlyBase; FBgn0026777; Rad23.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006636; STII.

```

```

DR InterPro; IPR000449; UBA domain.
DR InterPro; IPR000626; UbiQuitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiQuitin; 1.
DR SMART; SM00727; STII; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBQ; 1.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS0053; UBIQUITIN 2; 1.
SQ SEQUENCE 414 AA; 45780 MW; 4C2B494CA116F7AB CRC64;

Query Match 25.5%; Score 487.5; DB 5; Length 414;
Best Local Similarity 28.8%; Pred. No. 1.5e-26;
Matches 128; Conservative 78; Mismatches 130; Indels 109; Gaps 11;

Qy 1 MKLTVKTLKGTHTFEIRVQPNDDTMAVKNIETEEQKDSYFWGQQLLIFNGKVLKDESTLE 60
Db 1 MIITIKNLQOQTFTIEPAEKTIVLEAKKIFEEKGPE-YVAEKOKLIYAGVILTDRTVG 59
Qy 61 ENKVNEDGFLVWMLSKGKTSGTSTSSOHSN-----TPATROAPPLEAPQQAPOP----- 111
Db 60 SYNVDKFKFVWMLTRDSSSNRNQLSVKESNKLKLTSTDDSKQKMPCEANHTNSPSSNT 119
Qy 112 -----PVAPITTSQPEGLPAQAPNTHDNAASNLLSGRNVDITIINOLMENGSGGWDKDK 164
Db 120 EDSVLSRETRPLSSDELIGELAA-SLQSRASENLLMGDEYNQTVLSWVENG-----YPREQ 175
Qy 165 VQRALRAAYNPERAVEYLYSGIPV-----TABIAPVIGGGANTTDRAPTGE 212
Db 176 VERAMAAAYNPERAVEYLYNGIPAEKGTFNRLNENSTPLSPGQPASATSASERSTE 235
Qy 213 AGLSGINTAPLDLPQGNASGAGGGLDFLNNPQFQAVREMYHTNPQILQPLWVE 272
Db 236 -----SNSDPFEFLRSQFQLMKRSLIYQNPHELLHAVLQQ 270
Qy 273 LSKQNPOILRLIENHDEFLQLNEPF-----EGG----- 302
Db 271 IGQTNPALQLISENQDAFLNMLNQPIDRESESGATVPVSNARIPSTLDNVDLFSFDLE 330
Qy 303 -----EGDFLQPEDEMPHAIS-----VTPEQEAIGRLSEMGFDRARV 342
Db 331 VATSAQRSAAAGTSAAHQSGSAADN-EDLEQPLGVSTIRLNQKDKAIERLKALGFPEALV 389
Qy 343 IEAFLACDRNEELAANYLLEHAGEE 367
Db 390 LQAYFACEKNEEQANFLSSSFDD 414

```

Search completed: December 17, 2003, 06:20:07
Job time : 38.6093 secs

Result No.	Score	Query Match	Length	ID	Description
1	1910	100.0	368	21	Maize Rad23 protein
2	1169.5	61.2	371	21	Arabidopsis thaliana
3	1115	58.4	368	21	Arabidopsis thaliana
4	1067.5	55.9	348	21	Arabidopsis thaliana
5	1014	53.1	345	21	Arabidopsis thaliana
6	931.5	48.8	378	21	Arabidopsis thaliana
7	931.5	48.8	378	21	Arabidopsis thaliana
8	900.5	47.1	405	21	Maize Rad23 protein
9	871.5	45.6	299	21	Arabidopsis thaliana

Isolated nucleic acid encoding maize RAD23 protein is used to modulate the levels of polypeptides in plant or in assays for identifying

Isolated nucleic acid encoding maize RAD23 protein is used to modulate the levels of polypeptides in plant or in assays for identifying

PT compounds that bind to and/or increase/decrease enzymatic activity of
XX catalytically active polypeptides -
XX Claim 11b; Page 78-79; 82pp; English.
CC The present sequence is the maize Rad23 protein #2. It is isolated from
CC a Zea mays cell line, B73 callus tissue regenerated five days after
CC transfer of the callus from medium containing auxin to a medium devoid
CC of exogenous auxin. The cDNA is deposited under the ATCC No: PTA-531.
CC Maize Rad23 DNA sequence operably linked to a promoter, can be used to
CC construct a recombinant expression cassette. This expression cassette
CC can be used to generate a dicot or monocot transgenic plant e.g., maize,
CC soybean, sunflower, sorghum, canola, wheat, etc. It can also be used to
CC modulate the levels of Rad23 polypeptide expression in a plant or in
CC assays to identify compounds, that bind to and/or modulate the enzymatic
CC activity of catalytically active polypeptides.
XX SQ Sequence 368 AA;

Query Match 100.0%; Score 1910; DB 21; Length 368;
Best Local Similarity 100.0%; Pred. No. 7.8e-147;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLTIVTKLKGTHFEIRVOPNDTIMAVKKNIEIQKDSYPWGQQLIENGKVLKDESTLE 60
DB 1 MKLTIVTKLKGTHFEIRVOPNDTIMAVKKNIEIQKDSYPWGQQLIENGKVLKDESTLE 60

QY 61 ENKVNEDGLVVLMSKGTSGTSSQHSNTPATROAPPLEAPQAPQPPVAPITTSQ 120
DB 61 ENKVNEDGLVVLMSKGTSGTSSQHSNTPATROAPPLEAPQAPQPPVAPITTSQ 120

QY 121 PEGLPAPQPNTHDAAASLLSGRNVDITINQLMENGSGSWDKVQRAIRAAYNPERAV 180
DB 121 PEGLPAPQPNTHDAAASLLSGRNVDITINQLMENGSGSWDKVQRAIRAAYNPERAV 180

QY 181 EYLYSGIPVTAEIPIGQGGANTTDRAPTGAGISGIPNTAPLDLFPQGSNAGGGAGG 240
DB 181 EYLYSGIPVTAEIPIGQGGANTTDRAPTGAGISGIPNTAPLDLFPQGSNAGGGAGG 240

QY 241 GPLDLFRNNPQAVREVMVHTNPQILOPMLVELSKONFQILRLIEENHDEFLLQLNPEFE 300
DB 241 GPLDLFRNNPQAVREVMVHTNPQILOPMLVELSKONFQILRLIEENHDEFLLQLNPEFE 300

QY 301 GEGDFLDQPEDEPHAI SVTPPEQAEIGRLSEMGFDRARVIEAFLACDRNEELAANYL 360
DB 301 GEGDFLDQPEDEPHAI SVTPPEQAEIGRLSEMGFDRARVIEAFLACDRNEELAANYL 360

QY 361 LEHAGEED 368
DB 361 LEHAGEED 368

RESULT 2
AAG17377
ID AAG17377 standard; Protein; 371 AA.
XX AC AAG17377;
XX 17-OCT-2000 (first entry)
DT Arabidopsis thaliana protein fragment SEQ ID NO: 18372.
XX Arabidopsis thaliana.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0130077.
PR 19-APR-1999; 99US-0130449.
PR 21-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.


```
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0147303.
PR 27-JUL-1999; 99US-0147304.
PR 28-JUL-1999; 99US-0147302.
PR 02-AUG-1999; 99US-0147192.
PR 02-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.

PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 61.2%; Score 1169.5; DB 21; Length 371;
Best Local Similarity 64.3%; Pred. No. 1.2e-86;
Matches 247; Conservative 43; Mismatches 65; Indels 29; Gaps 9;

QY 1 MKLTVKTLKTHFEIRVQPNDTIMAVKKNIEIRIQKDSYPMGQQLIFNGKVLKDESTLE 60
DB 1 MKLTVKTLKSHFEIRVLPFSDTIMAVKKNIEDSQKNDYPCQQLIHNGKVLKDETSLV 60
QY 61 ENKVNEDGLVVLMSKGTSGTSSSOHSN-----TPATQAPPLEAPQAPQP 112
DB 61 ENKVTBEGFLVVLMSKSGSAGQSVQTSVSQFVSATTSSTKPAAP--STTQSSPVP 118
QY 113 VAPITTSQPEGLPAQAPNTHDNAENLLSGRNVDTIINQMEMGGGWDKVKQRAALRAA 172
DB 119 ASPIPAQEQ---PAAQTDTYQAAATLYSGSSLEQMVQOIMEMGGGWDKVTTRALRAA 175
QY 173 YNNPERAVLYSGIPVTAEIAP-----IGQGANTTDRAPTGBAGLSGIPNTAPLDLF 227
DB 176 YNNPERAVDLYSGIPQTAEVAVPPEAQIAGSGA-----ADVAPA--SGGPNSSPLDLF 228
QY 228 POGASNAGGAGGGLDFLRNNPQAVREWVHTNPOLQMLVLSKONPQILRIEEN 287
DB 229 POETVAAGSGDLGTLEFLRNNNDQQLRTVMVHSNPQLQPMQLGKQNPQLRIQEN 288
QY 288 HDEFLLQLNEPPEG--GEGDFLDQPEDEMPHAI SVTPEEQAI GRLESMPDRARVIEA 345
DB 289 QAEFLQLVNEPYEGSDGEGDMFDQEQ-EMPHAINVTAEQAIQRLEANGFDRLVIEA 347
QY 346 FLACDRNBELAAANYLLEHAGE-ED 368
DB 348 FLACDRNBELAAANYLLENSGDFED 371

RESULT 3
AAG44341
ID AAG44341 standard; Protein; 368 AA.
XX
```

AC AAG44341;
XX 18-OCT-2000 (first entry)
DT
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55531.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 27-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.

```
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-016192.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 58.4%; Score 1115; DB 21; Length 368;
Best Local Similarity 62.8%; Pred. No. 3,26-82;
Matches 235; Conservative 45; Mismatches 82; Indels 12; Gaps 8;

Qy 1 M K L T V K T L G T H F I R V Q P N D T I N A V K N I E E I Q K D S Y P W G Q O L L I F N G K V L K D E S T L E 60
Db 1 M K L T V K T L G S H F I R V L P T D T I N A V K N I E D S Q S K D N Y P C G Q O L L I H N G K V L K D E T T L V 60

Qy 61 E N K Y N D Q F L V M L S K G T S G T S S Q - H S N T P A T R Q A P P L E A P Q A P P V A P I T T S 119
Db 61 E N K Y T E G F L V V M L S K S T A S S A P S T Q P T I T T S I T T L A A P - S T T Q S T A V P A S N S 119

Qy 120 Q P - E G L P A Q P A N T H N A S N L L S G R N V D T I N Q L M E M G G S W D K K V O R A L R A A Y N N P R 178
Db 120 T P V Q E Q T A Q S D T Y G Q A A S T L V S G S S I E Q V Q Q I M E M G G S W D K E T V T R A L R A A Y N N P R 179

Qy 179 A V E Y L Y S G I P T A E I - A V P I G G Q G A N T T D R A P T G E A G L S G I P N T A P L D L F P Q A S N A G G 237
Db 180 A V D Y L Y S G I P T V T I P A T N L S G V G S G R E L T A P P P - - - - S G P N S P L D L F P Q E A V S D A A G 235
```

```
Qy 238 A C G G P L D F L R N N P Q F O A V R E M V H T N P Q I L Q P M L V E L S K N P Q I L R L I E N H D B F L Q L L N E 297
Db 236 G D L G T L E F L R G N D Q F O O L R S M V N S P Q I L Q P M L Q E L K N P Q I L R L I E N H D B F L Q L L N E 295
Qy 298 P F E G G E G - D P L D P E D E M P H A I S V T P R E Q E A I G R L E S M G P R A R V I E A F L A C D R N E E L 355
Db 296 P Y E G S D G V D I F D P D Q - E M P H S V N V T P E E Q S I E R L E A M G F D R A I V I E A F L S C D R N E E L 354
Qy 356 A A N Y L L E H A G E - E D 368
Db 355 A A N Y L L E H S A D F E D 368

RESULT 4
AAG17378
ID AAG17378 standard; Protein; 348 AA.
XX
AC AAG17378;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 18373.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 23-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
```

PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144634.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 55.9%; Score 1067.5; DB 21; Length 348;
Best Local Similarity 62.9%; Pred No. 2,1e-78;
Matches 227; Conservative 41; Mismatches 64; Indels 29; Gaps 9;

```
QY      24 MAVKKNIIEIOGKDSYPMQQLIFNGKVLKDESTLEENKKNEDGFLVVMLSKGKTSST 83
Db      1 MAVKKNIIEIOGKDNYPCCQQLIHNGKVLKDETSILVENKVTEEGFLVVMLSKSKSGSA 60
QY      84 GTSSQSHN-----TPATQAPPLEAPQAPQPPVAPITTSQPEGLPDAQAPNTHNA 135
Db      61 GOASVQTSSVQFVSATTSSTKPAAP--STTQSSVPVPSPIPAQEQ---PAAQTDTYGA 115
QY     136 ASNLLSGRNVDTIINOLMEMGGSDKKQVQRALRAAYNNPERAVEYLYSGIPVTAETAV 195
Db     116 ASLVSGSLEQVQVQIMEMGGSDKQVTVTRALRAAYNNPERAVDLYSYGIPQTAETAV 175
QY     196 P-----IGCGGANTTDRAPTGAGLSGIPNTAPLDLPQGNAGGAGGAGGGLDLFRNRP 250
Db     176 PVPEAQIAGSGA----APVAPA--SGGPNSSPLDLFPQETVAAAGSGDLGLTLELRND 228
QY     251 QQAVREMYHTNPQIOLPMLVLSKONPOILRLIENHDEFLOLLNEPPEG--GSGDFLD 308
Db     229 QFQQLRTWVHSNPQIOLPMLQBLGKONPOLLRLIQENQAEFTQLVNEPYEGSDGSGDMFD 288
QY     309 QPEEDEMPHAISVTPPEQEAIGRLSMGPDRAVIEAFACDRNBELAANYLLEHAGE-E 367
Db     289 QPEQ-EMPHAINVTPAEQAIQRLGAMGFDRAIVIEAFACDRNBELAANYLLENSGDFE 347
QY     368 D 368
Db     348 D 348

RESULT 5
AAG44342
ID  AAG44342 standard; Protein; 345 AA.
AC  AAG44342;
XX
DT  18-OCT-2000 (first entry)
DE  Arabidopsis thaliana protein fragment SEQ ID NO: 55532.
KW  Protein identification; signal transduction pathway; metabolic pathway;
KW  hybridisation assay; genetic mapping; gene expression control; promoter;
KW  termination sequence.
XX
OS  Arabidopsis thaliana.
XX
PN  EP1033405-A2.
XX
PD  06-SEP-2000.
XX
PF  25-FEB-2000; 2000EP-0301439.
XX
XX  25-FEB-1999; 99US-0121825.
PR  05-MAR-1999; 99US-0123180.
PR  09-MAR-1999; 99US-0123548.
PR  23-MAR-1999; 99US-0125788.
PR  25-MAR-1999; 99US-0126264.
PR  29-MAR-1999; 99US-0126785.
PR  01-APR-1999; 99US-0127462.
PR  06-APR-1999; 99US-0128234.
PR  08-APR-1999; 99US-0128714.
PR  16-APR-1999; 99US-0129845.
PR  19-APR-1999; 99US-0130077.
PR  21-APR-1999; 99US-0130449.
PR  23-APR-1999; 99US-0130510.
PR  28-APR-1999; 99US-0130891.
PR  28-APR-1999; 99US-0131449.
PR  30-APR-1999; 99US-0132048.
PR  30-APR-1999; 99US-0132407.
PR  04-MAY-1999; 99US-0132484.
PR  05-MAY-1999; 99US-0132485.
PR  06-MAY-1999; 99US-0132486.
PR  06-MAY-1999; 99US-0132487.
PR  07-MAY-1999; 99US-0132863.
PR  11-MAY-1999; 99US-0134256.
PR  14-MAY-1999; 99US-0134218.
PR  14-MAY-1999; 99US-0134219.
PR  14-MAY-1999; 99US-0134221.
PR  14-MAY-1999; 99US-0134370.
PR  18-MAY-1999; 99US-0134768.
PR  19-MAY-1999; 99US-0134941.
PR  20-MAY-1999; 99US-0135124.
PR  21-MAY-1999; 99US-0135353.
PR  24-MAY-1999; 99US-0135629.
PR  25-MAY-1999; 99US-0136021.
PR  27-MAY-1999; 99US-0136392.
PR  28-MAY-1999; 99US-0136782.
PR  01-JUN-1999; 99US-0137222.
PR  01-JUN-1999; 99US-0137528.
PR  04-JUN-1999; 99US-0137502.
PR  07-JUN-1999; 99US-0137724.
PR  08-JUN-1999; 99US-0138094.
PR  10-JUN-1999; 99US-0138540.
PR  10-JUN-1999; 99US-0138847.
PR  14-JUN-1999; 99US-0139119.
PR  16-JUN-1999; 99US-0139452.
PR  16-JUN-1999; 99US-0139453.
PR  17-JUN-1999; 99US-0139492.
PR  18-JUN-1999; 99US-0139454.
PR  18-JUN-1999; 99US-0139455.
PR  18-JUN-1999; 99US-0139456.
PR  18-JUN-1999; 99US-0139457.
PR  18-JUN-1999; 99US-0139463.
PR  18-JUN-1999; 99US-0139750.
PR  21-JUN-1999; 99US-0139763.
PR  21-JUN-1999; 99US-0139817.
PR  22-JUN-1999; 99US-0139899.
PR  23-JUN-1999; 99US-0140353.
PR  23-JUN-1999; 99US-0140354.
PR  24-JUN-1999; 99US-0140695.
PR  28-JUN-1999; 99US-0140823.
PR  29-JUN-1999; 99US-0140991.
PR  30-JUN-1999; 99US-0141287.
PR  01-JUL-1999; 99US-0141842.
PR  01-JUL-1999; 99US-0142154.
PR  02-JUL-1999; 99US-0142055.
PR  06-JUL-1999; 99US-0142390.
PR  08-JUL-1999; 99US-0142803.
PR  09-JUL-1999; 99US-0142920.
PR  12-JUL-1999; 99US-0142977.
PR  13-JUL-1999; 99US-0143542.
PR  14-JUL-1999; 99US-0143624.
PR  15-JUL-1999; 99US-0144005.
PR  16-JUL-1999; 99US-0144085.
PR  16-JUL-1999; 99US-0144086.
PR  19-JUL-1999; 99US-0144325.
PR  19-JUL-1999; 99US-0144331.
PR  19-JUL-1999; 99US-0144332.
PR  19-JUL-1999; 99US-0144333.
PR  19-JUL-1999; 99US-0144334.
PR  19-JUL-1999; 99US-0144335.
PR  20-JUL-1999; 99US-0144352.
PR  20-JUL-1999; 99US-0144632.
PR  20-JUL-1999; 99US-0144884.
PR  21-JUL-1999; 99US-0144814.
PR  21-JUL-1999; 99US-0145086.
PR  21-JUL-1999; 99US-0145088.
PR  22-JUL-1999; 99US-0145087.
PR  22-JUL-1999; 99US-0145089.
PR  22-JUL-1999; 99US-0145192.
```

Wed Dec 17 06:49:41 2003

PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 27-JUL-1999; 99US-0145319.
PR 27-JUL-1999; 99US-0145321.
PR 28-JUL-1999; 99US-0145386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159299.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.

PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161362.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 53.1%; Score 1014; DB 21; Length 345;
Best Local Similarity 61.3%; Pred. No. 4.7e-74;
Matches 215; Conservative 44; Mismatches 80; Indels 12; Gaps 8;

QY 24 MAVKKNIEIQKDSYPMGQQLLIIFNGKVLKDESTLEENKVNEDGLVVMLSKGTSGT 83
Db 1 MAVKKNIEDSQSKDNYPCGQQLLIHNGKVLKDETLVENKVTETGFLVVMLSKSTASSA 60
QY 84 GTSSSQ-HSNTPATRAPPEAPQAPQPPVAPITTSQP-EGLPAPAPNTHDNASNLIS 141
Db 61 GFSSTQPTSTTSTISSITTLAAP-STTQSIAPASNSTEVQOPTAQSDTYGQAATLVS 119
QY 142 GRNVDTIINQMEMGGGSHDKVQALRAAAYNNPERAVEYLYSGIPVTAET-AVPIGQ 200
Db 120 GSSIEQMVQIIMEMGGGSKETVTRALRAAYNNPERAVDYLYSGIPETVITPATNLGV 179
QY 201 GANTTDRAPTGEAGLSGIENPTAPLDLFPQASNAGGAGGGLDFLRNNPQFOAVREMH 260
Db 180 GSGRELTAPP-----SGGNSPFLDFPQENSDAAGGLTLEFLRGNDQFOQLRMVN 235
QY 261 TNPQILQPMVLVELSKONQILRLIENHDEFLLQNLNPEPFGEG--DFLDQPEDEMPHA 318
Db 236 SNFQILQPMQLQELGKQNPOLLRIQENQAEFLQLLNEPYEGSDGVDIFDQPDQ-EMPHS 294
QY 319 ISVTPEQEAIGRLSMGDFDRARVIEAFACDNEELAANYLLEHAGE-ED 368
Db 295 VNVTPQEESIERLEAMGDFRAIVIEAFILSCDRNEELAANYLLEHAGEADFED 345

RESULT 6
AAG19974
ID AAG19974 standard; Protein; 378 AA.
XX AAG19974;
AC AAG19974;
XX 17-OCT-2000 (first entry)
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 21981.
XX Arabidopsis thaliana.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
PD 25-FEB-2000; 2000EP-0301439.
PF 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134318.
PR 14-MAY-1999; 99US-0134319.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135829.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.

PR	13-OCT-1999;	99US-0159295.	
PR	14-OCT-1999;	99US-0159329.	
PR	14-OCT-1999;	99US-0159330.	
PR	14-OCT-1999;	99US-0159331.	
PR	14-OCT-1999;	99US-0159637.	
PR	14-OCT-1999;	99US-0159638.	
PR	18-OCT-1999;	99US-0159584.	
PR	21-OCT-1999;	99US-0160741.	
PR	21-OCT-1999;	99US-0160767.	
PR	21-OCT-1999;	99US-0160768.	
PR	21-OCT-1999;	99US-0160770.	
PR	21-OCT-1999;	99US-0160814.	
PR	21-OCT-1999;	99US-0160815.	
PR	22-OCT-1999;	99US-0160980.	
PR	22-OCT-1999;	99US-0160981.	
PR	22-OCT-1999;	99US-0160989.	
PR	25-OCT-1999;	99US-0161404.	
PR	25-OCT-1999;	99US-0161405.	
PR	25-OCT-1999;	99US-0161406.	
PR	26-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	
PR	26-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161920.	
PR	28-OCT-1999;	99US-0161992.	
PR	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	
Query Match 48.8%; Score 931.5; DB 21; Length 378;			
Best Local Similarity 52.6%; Pred. No. 2.7e-67;			
Matches 200; Conservative 56; Mismatches 109; Indels 15; Gaps 8;			
Qy	1	MKLTVKTLKGTHFIRVOPNDTMAVKNIIEIQKDSYPMGQQLLIFNGKVLKDESTLE	60
Db	1	MKIFVKTLSSGNFIEVKPADKVDKTAIETVKGAE-YPAAKQMLIHQGVKLVKDETTLE	59
Qy	61	ENKYNEDFLVMLSKGKTS---GSTGTSSOHSNTPATQAPLEAP-QOAPQPPVAPI	116
Db	60	ENNVENSFIVIMLSKTKASPSGASTASAPASATQCTVATPQVSAPTASVPVTSSTA	119
Qy	117	TTSOPEGLPAQAPNTHDNAAGNLLSGRNVDTIINQLEMGGGSKDKQVQALRAAANNP	176
Db	120	TAAAPATAASVQTDVYQGAASNLVAGTTLESTVQQLDMGGGSDRDTVVRAALRAAFNP	179
Qy	177	ERAVEYLSGIPVTAIEI-----AVPIGCGGAN--TTDRAPTGEAGLSGIPNTAPLDLFPQ	229
Db	180	ERAVEYLSGIPQAQAEIPFVAQAPATGEOANPLAQPOQAAAPAAATGCPNANPLNLPQ	239
Qy	230	GASNAGGAGGGLDFLRNPNQFQAVREVMVHTNPQILQPLMLVELSKQNPQILRIEENHD	289
Db	240	GMPAADAGAGNLDLFLNSQGFQALRAMVQANPQILQPLMLQELGKQNPQLVRLIQEHOA	299
Qy	290	EFLQLLNEFFEGGDFLDQPEEDMPHAI SVTPEEQAIGRLSMGDFDRARVTEAFIAC	349
Db	300	DFRLRLINEPVE-GEENVMEQLEA-AMPQAVTVPTEERAEIERLEGMGDFRAMVLEVFAC	357
Qy	350	DRNEELANLYLLEHAGE-ED 368	
Db	358	NKNEELANLYLIDHMHFEFD 377	
RESULT 7			
AAG45204			
ID	AAG45204 standard; Protein; 378 AA.		
XX			
AC	AAG45204;		
XX			
DT	18-OCT-2000 (first entry)		
XX			
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 56723.		
XX			
DE	Protein identification; signal transduction pathway; metabolic pathway;		
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		
KW	termination sequence.		

XX	OS	Arabidopsis thaliana.	
XX	PN	EP1033405-A2.	
XX	PD	06-SEP-2000.	
XX	XX	25-FEB-2000; 2000EP-0301439.	
XX	XX	25-FEB-1999;	99US-0121825.
PR	05-MAR-1999;	99US-0123180.	
PR	09-MAR-1999;	99US-0123548.	
PR	23-MAR-1999;	99US-0125788.	
PR	25-MAR-1999;	99US-0126264.	
PR	29-MAR-1999;	99US-0126785.	
PR	01-APR-1999;	99US-0127462.	
PR	06-APR-1999;	99US-0128234.	
PR	08-APR-1999;	99US-0128714.	
PR	16-APR-1999;	99US-0129845.	
PR	19-APR-1999;	99US-0130077.	
PR	21-APR-1999;	99US-0130449.	
PR	23-APR-1999;	99US-0130510.	
PR	23-APR-1999;	99US-0130891.	
PR	28-APR-1999;	99US-0131449.	
PR	30-APR-1999;	99US-0132048.	
PR	30-APR-1999;	99US-0132407.	
PR	04-MAY-1999;	99US-0132484.	
PR	05-MAY-1999;	99US-0132485.	
PR	06-MAY-1999;	99US-0132486.	
PR	06-MAY-1999;	99US-0132487.	
PR	07-MAY-1999;	99US-0132863.	
PR	11-MAY-1999;	99US-0134256.	
PR	14-MAY-1999;	99US-0134218.	
PR	14-MAY-1999;	99US-0134219.	
PR	14-MAY-1999;	99US-0134221.	
PR	14-MAY-1999;	99US-0134370.	
PR	18-MAY-1999;	99US-0134768.	
PR	19-MAY-1999;	99US-0134941.	
PR	20-MAY-1999;	99US-0135124.	
PR	21-MAY-1999;	99US-0135353.	
PR	24-MAY-1999;	99US-0135829.	
PR	25-MAY-1999;	99US-0136021.	
PR	27-MAY-1999;	99US-0136392.	
PR	28-MAY-1999;	99US-0136782.	
PR	01-JUN-1999;	99US-0137222.	
PR	03-JUN-1999;	99US-0137528.	
PR	04-JUN-1999;	99US-0137502.	
PR	07-JUN-1999;	99US-0137724.	
PR	08-JUN-1999;	99US-0138094.	
PR	10-JUN-1999;	99US-0138540.	
PR	10-JUN-1999;	99US-0138847.	
PR	14-JUN-1999;	99US-0139119.	
PR	16-JUN-1999;	99US-0139452.	
PR	16-JUN-1999;	99US-0139453.	
PR	17-JUN-1999;	99US-0139492.	
PR	18-JUN-1999;	99US-0139454.	
PR	18-JUN-1999;	99US-0139455.	
PR	18-JUN-1999;	99US-0139456.	
PR	18-JUN-1999;	99US-0139457.	
PR	18-JUN-1999;	99US-0139458.	
PR	18-JUN-1999;	99US-0139459.	
PR	18-JUN-1999;	99US-0139460.	
PR	18-JUN-1999;	99US-0139461.	
PR	18-JUN-1999;	99US-0139462.	
PR	18-JUN-1999;	99US-0139463.	
PR	18-JUN-1999;	99US-0139750.	
PR	18-JUN-1999;	99US-0139763.	
PR	21-JUN-1999;	99US-0139817.	
PR	22-JUN-1999;	99US-0139899.	
PR	23-JUN-1999;	99US-0140353.	
PR	23-JUN-1999;	99US-0140354.	
PR	24-JUN-1999;	99US-0140695.	
PR	28-JUN-1999;	99US-0140823.	

PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.

PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 22-SEP-1999; 99US-0155186.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 48.8%; Score 931.5; DB 21; Length 378;
Best Local Similarity 52.6%; Pred No. 2.7e-67;

Matches 200; Conservative 56; Mismatches 109; Indels 15; Gaps 8;

Qy 1 MKLTVKTLKGTHEIRVQPNDDTMAVKKIEBIOGKDSYPMGQQLIFNGKVLKDESTLE 60
Db 1 MKIEFKTLGSGNFEIEVKPADKVDVKTATETVKGAE-YPAAKQMLIHQKVLKDESTLE 59
Qy 61 ENKYNEDGFLVWMLSKGKTS---GSTGTSSQHSNTPATQAAPLEAP-QAOPPPVAPI 116
Db 60 ENNVVENSFVIMLSKTKASPGASTASAPASATQPTVATPQVSAPTASVPVPTSGTA 119
Qy 117 TTISOEGLPAQAPNTHDNAASNLISGRNVDTIINQLEMGGSGWDKDVORALRAAYNPP 176
Db 120 TAAAPATAASVQTDVYCGAASNLVAGTTLESTVQOILDMGGSGWDRDTVVRLRAAFNPP 179
Qy 177 ERAVEYLYSGIPVTAETI-----AVPIGGQGAN--TTDRAPTGEAGLSGIPNTAPLDLFPQ 229
Db 180 ERAVEYLYSGIPVTAETI-----AVPIGGQGAN--TTDRAPTGEAGLSGIPNTAPLDLFPQ 239
Qy 230 GASNAGGAGGGPLDFLRNPPQAVREMHVHTNPQIQLQPMVLVELSKONPQILRLIENHND 289
Db 240 GMPAADAGAGAGNLDFLRNSQQFQALRAMVQANPQILQPMQLQELGKQNPQLVRLIQEHOA 299
Qy 290 EFLQLLNEPPEGSGDFLDQPEDEMEPHAISVTPPEOEAIERLESMDGFRARVIEAFAC 349
Db 300 DFLRLINEPVE-GEENVMEQLEA-AMPQAVTVTPPEEREATERLEGMGDFRAMVLEVPFAC 357
Qy 350 DRNEELAANYLLEHAGE-ED 368

PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149446.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match
Best Local Similarity

45.6%; Score 871.5; DB 21; Length 299;
60.3%; Pred. No. 1.5e-62;

Matches 188; Conservative 36; Mismatches 59; Indels 29; Gaps 9	
QY	73 MLSKGTSGTGTSSQHSN-----TPATRAQPPLEAQQOQPQPVPATITTSQPEGL 124
Db	1 MLSKSGSGSAGQASVQTSSVQPSVATTSSTKPAAP--STTQSSVPASPAPAQEQ-- 55
QY	125 PQAQPNTHDNAASNULLSGRNVDTTIINQMEMGGGSHDKVQRALRAAYNNPERAVEVLY 184
Db	56 PAAQDTDTGYAASTVSGSSLEQMVQOQIMEMGGGSHDKETVTRALRAAYNNPERAVDLY 115
QY	185 SGIPVTAETAVP-----IGGOGANTTDRAPTGEBAGLSGIPNTAPLDLFPQGASNAGGCGAG 239
Db	116 SGIPQTAETAVVPVEAQIAGSGA-----APVAPA--SGGNSSPDLDFQETVAAGSGD 168
QY	240 GGLDFLRNNPQFOAVREMVHTNPQIILQPMVLVELSKONPQILRLIENHDEFQLLNPPF 299
Db	169 LCTLEFLRNNDFQQLRTMVHSNPQIILQPMQLQELGKNPQLRLIQENQAQEFQLVNEPY 228
QY	300 EG--GEGDFLQPEDEENPHAI SVTPSEQBAIGRLSMGFDPRARVTEAFACDRNEELAA 357
Db	229 EGSGEGEGWDFQPEQ-EMPHAINVTPAEQBAIQRLAMGFDPRALVTEAFACDRNEELAA 287
QY	358 NYLLEHAGE-ED 368
Db	288 NYLLENSGDFED 299
RESULT 10	
AAG19975	
ID	AAG19975 standard; Protein; 335 AA.
XX	AAG19975;
DT	17-OCT-2000 (first entry)
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 21982.
DE	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	Arabidopsis thaliana.
OS	Arabidopsis thaliana.
XX	EP1033405-A2.
XX	06-SEP-2000.
XX	25-FEB-2000; 2000EP-0301439.
PF	25-FEB-1999; 99US-0121825.
XX	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	23-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130449.
PR	23-APR-1999; 99US-0130510.
PR	23-APR-1999; 99US-0130891.
PR	28-APR-1999; 99US-0131449.
PR	30-APR-1999; 99US-0132048.
PR	30-APR-1999; 99US-0132407.
PR	04-MAY-1999; 99US-0132484.
PR	05-MAY-1999; 99US-0132485.
PR	06-MAY-1999; 99US-0132486.
PR	06-MAY-1999; 99US-0132487.
PR	07-MAY-1999; 99US-0132863.
PR	11-MAY-1999; 99US-0134256.
PR	14-MAY-1999; 99US-0134218.

```
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0160989.

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 43.7%; Score 834; DB 21; Length 335;
Best Local Similarity 53.3%; Pred. No. 1.9e-59;
Matches 179; Conservative 47; Mismatches 96; Indels 14; Gaps 7;

Qy 45 LLIFNGKVLDESTEENKVNEDFLVVLKSGKTS---GSTGTSSSOHSNTPATROAPP 101
Db 1 MLIHGKVLKDETTLENNVVENSVFVIMLSKTKASPSGASTASAPASATQPPQVATPQ 60

Qy 102 LEAP-QQAPQPPVAPITTSQPEGLPAQAPNTHDNAASNLISGRNVDTIINOLMENGGSW 160
Db 61 VSAPTASVPVPTSGTATAAATAASVQTDVYQASNLVAGTTLSTVQIILDMGGGSW 120

Qy 161 DKDKVQRALRAAYNNPERAVEYLISGIPVTABI-----AVPIGGQGAN--TTDRAPTGEA 213
Db 121 DRDVTVRALRAAFNNPERAVEYLISGIPQAQAEIPPAQAPATGEGQAANPLAQPOQAAAPA 180

Qy 214 GLSGIPNTAPLDLPQGNASNAGGGGGLDPLRNPNQFOAVRENVHTNPQILOPMLVEL 273
Db 181 AATGGPNANPLNLPQGMPAADAGAGNLDPLRNSQOQALRAVAVQANPQILOPMLQEL 240

Qy 274 SKQNPQLRLJEENHDEFLQLNPEFEGGDFLDQPEEDSMPHAISVTPPEQEAIGRLE 333
Db 241 GKQNPQLVRLIQEHQADFLRLINEPVE-GEENVMEQLEA-AMPQAVTVPPEEAIEGLE 298

Qy 334 SMGFDRARVIEAFLACDRNEELAANYLLEHAGE-ED 368
Db 299 GMGFDRAMVLEVFACNKEELAANYLLDHMHPEFD 334

RESULT 11
AAG45205
ID AAG45205 standard; Protein; 335 AA.
AC AAG45205;
XX
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 56724.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
```

PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142800.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.

PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match		43.7%;	Score 834;	DB 21;	Length 335;
Best Local Similarity		53.3%;	Pred. No. 1.9e-59;		
Matches 179;		Conservative 47;	Mismatches 96;	Indels 14;	Gaps 7

Qy	45	LLIFNGKVLKDESLTEENKYNVEDGLVWMLSGKTS----	GSTGTSSSOHSNTPATRQAPP	101
Db	1	MLIHQGVKLDTELTLEENNVVENSEFIVMLSKTKASP	GASTASAPSATQPQTVA	60
Qy	102	LEAP--QQAQPPVPAPITTSQPEGLPAQAPNTHDNAAS	NLLSGRNVDTIINOLMEMGGGSW	160
Db	61	VSAPTASVPVTSGTATATAATAASVQTDVYQAA	SNLVAGTTLSTVQQLLDMGGGSW	120
Qy	161	DKDKVQRALRAAYNNPRAVELYSGIPVTABI-----	AVPIGGQGAN--TTDRAPTGEA	213
Db	121	DRDTVVRALRAAFNNPRAVELYSGIPAAAEIP	PPVAQAPATGEAAANPLAQ	180
Qy	214	GLSGIPNTAPLDLFPQGNASNAGGAGGGLDFLR	NNPOFOAVREMVHTNPOILQ	273
Db	181	AATGGPNANPLNLPQCGMPAADAGAGAGNLDF	LRNSQOFOALRAMVQANPOT	240
Qy	274	SKQNPQILRLTEENHDEFLOLLNPEFEGEGD	FLDQPEEDSMPHAISVTP	333
Db	241	GKQNPQLVRLTQEQHADFLRLINEPVE--GE	ENVMEQLEA--AMPQAVTV	298
Qy	334	SMGFDRARVIEAFACDRNEELAAANYLLEHAGE	--ED 368	
Db	299	GMGFDRAMVLVEVFACNKNNEELAAANYLLD	HMHEFED 334	

RESULT 12	
AAG44343	
ID	AAG44343 standard; Protein; 296 AA.
XX	
AC	AAG44343;
XX	
DT	18-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 55533.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	

PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144325.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	12-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	13-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match 42.9%; Score 820; DB 21; Length 296;		
Best Local Similarity 58.3%; Pred. No. 2.2e-58;		
Matches 176; Conservative 40; Mismatches 74; Indels 12; Gaps 8;		
QY	73	MLSKGKTSSTGTSSQ-HSNTPATRQAPPLPAPQAPQPPVAPITTSQP-EGLPQAQPN 130
Db	1	MLSKSKTASSAGPSSTQPTSTTSTTSLAAP-STTQSIAPASNSTPVQRPQTAQSD 59
QY	131	THDNAASNLGSRNVDITINQLMEMGGGSKDKVORALRAAYNNPERAVEVLYSGIPVT 190
Db	60	TYGQAASLTLYSGSSIEQMVGQIQIMEMGGGSKDKETVTRALRAAYNNPERAVDYISGIPET 119
QY	191	AEI-AVPIGGOGANTTDRAPTGEAGLSGIPNTAPLDLPQAGSNAGGAGGGLDLFRNN 249
Db	120	VTIPATNLGSGVSGRELTAAPP-----SGGPNSSPLDLFPQEAVSDAAGGDLGTLEFLRGN 175
QY	250	PFQAVREVMYHTNPQILQPMVLKSNQKQILRIENHDEFLQLLNEPPEGEG--DFL 307
Db	176	DQFQQLRSMVSNPQILQPMVLKSNQKQILRIENHDEFLQLLNEPPEGEG--DFL 235
QY	308	DQPEEEMPHAI SVTPEEQEAI GRLSMGDFDRARVEAFACDRNEELAANYLLEHAGE- 366
Db	236	DQPDQ-EMPHSVNVTPPEEQESIERLEAMGDFRAIVIEAFLSCDRNEELAANYLLEHAGE- 294
QY	367	ED 368
Db	295	ED 296
RESULT 13		
AAG19976		
ID	AAG19976 standard; Protein; 307 AA.	
XX		
AC	AAG19976;	
XX		
DT	17-OCT-2000 (first entry)	

PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match 38.8%; Score 742; DB 21; Length 307;		
Best Local Similarity 52.3%; Pred. No. 5.1e-52;		
Matches 161; Conservative 42; Mismatches 91; Indels 14; Gaps 7;		
Qy	73	MLSGKTS---GSGTSSSHSNTPATROAPPLEAP-QQAPQPPVAPITTSQEGPLPAQA 128
Db	1	MLSKTKASPGSASTASAPASATOPQTATPQVSAPTASVEPTSGTATAAPATAASVQ 60
Qy	129	PNTHDNAANLLSRNVDTIINQMENGSGSDKQYORALRAAYNNPERAVEVLYSGIP 188
Db	61	TDVYQAAANLVAGTTLETTSVQQLDMGGGSDRDTTVRALRAAFNNPERAVEVLYSGIP 120
Qy	189	VTAEI-----AVPTGGGAGN--TTDRAPTGEAGLSGIPNTAPLDLFPQGSNAGGGAGGG 241
Db	121	AQAEIPVQAQAPATGEQANPLAQQAAPAAATGFPNANPLNLFQGMFADAGAGAG 180
Qy	242	PLDFLRNNPQAVREVHTNPQILQPMVLVELSKNQFQILRLIEENHDFLQLNEPFG 301
Db	181	NLDFLRNSQQQALRAVQANPQILQPMVLQELGKQNPQLVRIQEQHQADEFRLINEPVE- 239
Qy	302	GEGDFLOQPEDEPHAISVTPPEQEAIGRLSNGFDRARVIEAFACDRNEELAANYLL 361
Db	240	GEENVMEQLEA-AMFQAVTVPTEAREATLERLGGGFDRAVLYEVFFACNKEELAANYLL 298

Qy	362	EHAGE-ED 368
Db	299	DHMHEPED 306
RESULT 14		
RAG45206		
IX	AAG45206 standard; Protein; 307 AA.	
XX	AAG45206;	
DT	18-OCT-2000 (first entry)	
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 56725.	
DE	Arabidopsis thaliana	
KW	Protein identification; signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
KW	termination sequence.	
OS	Arabidopsis thaliana.	
XX	EP1033405-A2.	
PD	06-SEP-2000.	
XX	25-FEB-2000; 2000EP-0301439.	
PR	25-FEB-1999;	99US-0121825.
PR	05-MAR-1999;	99US-0123180.
PR	09-MAR-1999;	99US-0123548.
PR	23-MAR-1999;	99US-0125788.
PR	25-MAR-1999;	99US-0126264.
PR	29-MAR-1999;	99US-0126785.
PR	01-APR-1999;	99US-0127462.
PR	06-APR-1999;	99US-0128234.
PR	08-APR-1999;	99US-0128714.
PR	16-APR-1999;	99US-0129845.
PR	19-APR-1999;	99US-0130077.
PR	21-APR-1999;	99US-0130449.
PR	23-APR-1999;	99US-0130510.
PR	23-APR-1999;	99US-0130891.
PR	28-APR-1999;	99US-0131449.
PR	30-APR-1999;	99US-0132048.
PR	30-APR-1999;	99US-0132407.
PR	04-MAY-1999;	99US-0132484.
PR	05-MAY-1999;	99US-0132485.
PR	06-MAY-1999;	99US-0132486.
PR	06-MAY-1999;	99US-0132487.
PR	07-MAY-1999;	99US-0132863.
PR	11-MAY-1999;	99US-0134256.
PR	14-MAY-1999;	99US-0134218.
PR	14-MAY-1999;	99US-0134219.
PR	14-MAY-1999;	99US-0134221.
PR	14-MAY-1999;	99US-0134370.
PR	18-MAY-1999;	99US-0134768.
PR	19-MAY-1999;	99US-0134941.
PR	20-MAY-1999;	99US-0135124.
PR	21-MAY-1999;	99US-0135353.
PR	24-MAY-1999;	99US-0135629.
PR	25-MAY-1999;	99US-0136021.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.

Db 61 TDVYGOASNLVAGTTLESTVQIILDMGGSHDRDVTIVRALRAAFNNPRAVEYLYSGIP 120
QY 189 VTAEI-----AVPIGQGAN--TTDRAPTGEAGLSGIPNTAPLDLFPQASNAGGAGGG 241
Db 121 AQAEIPVQAAPATGEOANPLAQPOAAAAPAAATGGPNANPLNLPQGMPAADAGAGAG 180
QY 242 PLDFLNNPQFQAVREWHNTNPOILQPLMLVELSKNPQILRLTEENHDSFLOLLNEPPRG 301
Db 181 NLDFLNSOQFQALRAMVQANPQILQPLMLQELGKNPQVRLVLIQHQADFRLINEPVB- 239
QY 302 GEGDFLDQPEEDMPHAI SVTPPEQEAIGRLSMGFDRAVIEAFIACDRNEBELAANYLL 361
Db 240 GEENVWEQLEA-AMPQAVTVTPPEREAIERLEGMGFDRAWLVEVFFACNKNEELANYLL 298
QY 362 EHAGE-ED 368
Db 299 DHMHEPED 306

RESULT 15
ABB57171
ID ABB57171 standard; Protein; 416 AA.
XX
AC ABB57171;
XX
DT 07-MAR-2002 (first entry)
XX
DE Mouse ischaemic condition related protein sequence SEQ ID NO:415.
XX
KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease.
XX
OS Mus musculus.
XX
XX W0200188188-A2.
XX
XX 22-NOV-2001.
XX
XX 18-MAY-2001; 2001WO-JP04192.
XX
XX 18-MAY-2000; 2000JP-0145977.
XX
XX (UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
XX Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
XX WPI; 2002-034733/04.
XX
XX N-PSDB; ABI99464.

Examining the ischemic condition (e.g. occlusive ischemia) by measuring expression levels of particular genes defined in the specification or by determining the expression profile of a gene group comprising these genes -
Claim 2; Page 1132-1134; 2690pp; English.
The present invention describes a method for examining ischaemic conditions, comprising measuring the expression levels of particular genes (I) in a test sample or determining the expression profile of a gene group in the sample comprising genes selected from (I). The method is useful for examining the ischaemic condition (e.g. compressive ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring expression levels of particular genes (ABI99202 to ABI99912, encoding the protein sequences in ABB57020 to ABB57374) or by determining the expression profile of a gene group comprising these genes. The expression levels or expression profiles produced by these genes are used as an indicator when screening for ischaemic condition-improving drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR primers for a mouse ischaemic condition related sequence, which are used in the exemplification of the present invention.

XX Sequence 416 AA;

Query Match 33.2%; Score 634; DB 23; Length 416;
Best Local Similarity 33.5%; Pred. No. 4.6e-43;
Matches 146; Conservative 93; Mismatches 109; Indels 88; Gaps 11;
QY 1 MKLVTTLKGTHTFEIRVOPNDTIMAVKNIEIIOCKDSYPWGQQLLIIFNGKVLKDESTLE 60
Db 1 MQVTLKTLQQOQTFKIDIDPEETVKALKKEIKSEKGDAPPVAGQKLIYAGKILSDDTALK 60
QY 61 ENKYNEDFLVVMLSKGTSGTGTSSSHQHNTPA---TROAPPLEAPQ-OAQPQPPVAPI 116
Db 61 EYKIDKFNVVVMVTKAVTAVPATTPQSTSTPSTTVSSSPAVAAAQAAPTPALAPT 120
QY 117 TT-----SQPE---GLPAQAP-----NTHONA 135
Db 121 STPATSTTASSEPAPAGATQPKPAEKPAQTFVLTSAPADSTPCDSSRSNLFEDA 180
QY 136 ASNLLSGRNVDTIINQLMEMGGGSHDKVORALRAAYNNPERAVEYLYSGIPVTAETAV 195
Db 181 TSALVTGQSYENMVTEIMSMG---YEREQVIAALRAFNNPDRAVEYLLMGIPGDRSQQA 237
QY 196 PIGGOGANTTDRAPTGEAGLSGIPNTAPLDLFPQASNAG-----GGAGGGPLDFLRN 248
Db 238 VVD-----PPQAVSTGTPOS-----PAVAAAAATTTATTTTSGHPLFLRN 281
QY 249 NPQFOAVREWHNTNPOILQPLMLVELSKNPQILRLIEENHDEFLOLLNEPPE- 300
Db 282 QPQFOQMRQIIQONPSLLPALLOQIGRENPNLLOQISOHQEHFIQMLNEPVOEAGGQGGG 341
QY 301 -----GGEGDFLDQPEEDMPHAI SVTPPEQEAIGRLSMGFDRAVIEAFIACDRN 352
Db 342 GGGGGGGGGGGGGGIAEAGSGHMNY-IQVTPQEKAEIRLKGALGPFPEGLVIOAYFACEKN 400
QY 353 BELAANYLLEHAGEED 368
Db 401 ENLAANFLQQNFDED 416

Search completed: December 17, 2003, 06:18:39
Job time : 36.229 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 06:20:11 ; Search time 126.158 Seconds
(without alignments)
544.771 Million cell updates/sec

Title: US-09-805-550-4

Perfect score: 1910

Sequence: 1 MKLTVKTLKGTHTFIRVQPN.....CDRNEELANVLEHAGEED 368

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 696363 seqs, 186758610 residues

Total number of hits satisfying chosen parameters: 696363

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pdp.*
12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pdp.*
13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1910	100.0	368	9	US-09-805-550-4
2	900.5	47.1	405	9	US-09-805-550-2
3	603.5	31.6	363	12	US-10-116-275-170
4	603	31.6	379	15	US-10-102-806-549
5	182	9.5	79	11	US-09-918-036-4
6	176	9.2	589	12	US-10-293-000-4
7	176	9.2	595	12	US-10-293-000-2
8	164	8.6	81	11	US-09-918-036-5
9	150	8.1	522	15	US-10-296-770-2
10	150	7.9	624	15	US-10-146-473-47
11	140	7.3	504	11	US-09-918-864-410
12	140	7.3	504	12	US-10-195-835-410
13	140	7.3	504	12	US-10-286-333-410
14	140	7.3	504	15	US-10-125-635A-410
15	140	7.3	504	15	US-10-002-603-410

16 139 7.3 63 9 US-09-925-299-1544
17 139 7.3 63 11 US-09-925-299-1544
18 134 7.0 421 15 US-10-296-770-7
19 133.5 7.0 1274 15 US-10-020-215-2
20 131 6.9 139 12 US-10-263-828-120
21 130 6.8 966 12 US-10-304-454-2
22 129 6.8 655 14 US-10-001-632A-2
23 129 6.8 655 15 US-10-054-683-29
24 128.5 6.7 336 10 US-09-987-107-44
25 128.5 6.7 390 15 US-10-102-806-461
26 128 6.7 337 10 US-09-987-107-46
27 127 6.6 388 12 US-10-094-749-2859
28 126.5 6.6 160 10 US-09-738-973-205
29 126.5 6.6 160 10 US-09-854-133-205
30 126.5 6.6 160 15 US-10-144-649A-205
31 126.5 6.6 228 12 US-10-029-386-33570
32 124.5 6.5 710 9 US-10-296-770-4
33 123.5 6.5 110 15 US-09-925-299-1042
34 123.5 6.5 110 11 US-09-925-299-1042
35 123 6.4 151 12 US-10-360-053-10
36 122.5 6.4 76 9 US-09-804-866-16
37 122.5 6.4 76 9 US-09-826-312-13
38 122.5 6.4 76 11 US-09-918-036-1
39 122.5 6.4 76 14 US-10-058-820-21
40 122.5 6.4 76 15 US-10-108-767-13
41 122.5 6.4 76 15 US-10-152-156-13
42 122.5 6.4 76 15 US-10-275-985-1
43 122.5 6.4 128 12 US-10-205-194-87
44 122.5 6.4 159 12 US-10-360-053-12
45 122.5 6.4 530 15 US-10-296-770-5

ALIGNMENTS

RESULT 1

US-09-805-550-4
; Sequence 4, Application US/09805550
; Patent No. US20020026045A1
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; FILE REFERENCE: Tagliani, Laura
; TITLE OF INVENTION: Rad23 Genes and Uses Thereof
; CURRENT APPLICATION NUMBER: US/09/805,550
; CURRENT FILING DATE: 2001-03-13
; PRIOR FILING DATE: 09/413,574
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: 60/109,728
; PRIOR FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Zea mays
US-09-805-550-4

Query Match 100.0%; Score 1910; DB 9; Length 368;
Best Local Similarity 100.0%; Pred. No. 5.6e-148;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MKLTVKTLKGTHTFIRVQPNNDTMAVKKNIEETQKDSYPMGQOLLIFNGKVLKDESTLE 60
Db 1 MKLTVKTLKGTHTFIRVQPNNDTMAVKKNIEETQKDSYPMGQOLLIFNGKVLKDESTLE 60
Qy 61 ENKVNEDGFLVNLMSKGTSGTGTSSQHSNTPATRAQAPPLAPQAPPPVAPITTSQ 120
Db 61 ENKVNEDGFLVNLMSKGTSGTGTSSQHSNTPATRAQAPPLAPQAPPPVAPITTSQ 120
Qy 121 PEGLPQAQPNTHNAAANLLSGRNVDTIINQLMEMGGSGWDKQVQALRAAYNNPERAV 180
Db 121 PEGLPQAQPNTHNAAANLLSGRNVDTIINQLMEMGGSGWDKQVQALRAAYNNPERAV 180

QY 181 EYLSGIPVTABIAVPIGGQANTTDRAPTGAGLSGIPNTAPLDLFPQASNAGGAGG 240
Db 181 EYLSGIPVTABIAVPIGGQANTTDRAPTGAGLSGIPNTAPLDLFPQASNAGGAGG 240
QY 241 GPLDLRNNPQAVREVMHTNPQILQPMVLSTKONPQILRLIENHDEFLLNNEPFE 300
Db 241 GPLDLRNNPQAVREVMHTNPQILQPMVLSTKONPQILRLIENHDEFLLNNEPFE 300
QY 301 GEGDFLOPDEEMPHATSVTPEQEAIGRLSMGDFRARIETAFACDRNEELAAAYL 360
Db 301 GEGDFLOPDEEMPHATSVTPEQEAIGRLSMGDFRARIETAFACDRNEELAAAYL 360
QY 361 LEHAGEED 368
Db 361 LEHAGEED 368

RESULT 2
US-09-805-550-2
; Sequence 2, Application US/09805550
; Patent No. US2002026045A1
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Rad23 Genes and Uses Thereof
; FILE REFERENCE: 0964D
; CURRENT APPLICATION NUMBER: US/09/805,550
; CURRENT FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 09/413,574
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: 60/109,728
; PRIOR FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Zea mays
US-09-805-550-2

Query Match 47.1%; Score 900.5; DB 9; Length 405;
Best Local Similarity 49.0%; Pred. No. 2.4e-65;
Matches 197; Conservative 53; Mismatches 117; Indels 35; Gaps 7;
QY 1 MCLTVKTLKGTHTFEIRVQPNNDTMAVKNIIEEIQKDSYPWGQQLIFNGKVLKDESTLE 60
Db 1 MCLNVKTLKGTHTFEIRVQPNNDTMAVKNIIEEIQKDSYPWGQQLIFNGKVLKDESTLE 60
QY 61 ENKYNEDGFLVWMLSKGTSSTSSQHSNTPATRQAP-----PLEAPQAP-QPPVA 114
Db 61 SNGVAENSLVIMLSKAKAS-SSGASTATTAKAPATLAQPAAPVAPASVARTPTQAPVA 119
QY 115 PITTSQPEGLPAQAP-----NTHDNAASLLSGRNVDTIINLMEMGGSDKDK 164
Db 120 TAETAPPSVQQAAPATAVATDADVYQAASNLVFNQNNLEQTIQILDGNGGTWERDT 179
QY 165 VQALRAAYNNPERAVEYLISGIPVTAE-----IAVPIGGQANTTDRAPTGEA----- 213
Db 180 VVRLAAYNNPERAIDYLSGIPENVEAQPVAPAAQGOOTNOOAAAPQAPVALPVQP 239
QY 214 -GLSGIPNTAPLDLFPQASNAGG-----GAGGGLDLRNNPQFQAVREVMHTNPQIL 266
Db 240 SPASAGPNNPNNLFPQVPSGGNPGVVPAGSGALDALRQLPQFQALLQVLQANPQIL 299
QY 267 QPMLVELSKONPQILRLIENHDEFLLNNEPFEFGEGDFLDQPEEDEMPHAISVTPEEQ 326
Db 300 QPMLQELKQNPQILRLIQENQAEBFLRVNESPESGPGCNILGQLAAAVPQTLVTPTEER 359
QY 327 EAIQRLSMGDFRARIETAFACDRNEELAAAYLLEHAGEED 368
Db 360 EAIQRLSMGDFRARIETAFACDRNEELAAAYLLEHAGEED 401

RESULT 3
US-10-116-275-170
; Sequence 170, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elan Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 170
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-275-170

Query Match 31.6%; Score 603.5; DB 12; Length 363;
Best Local Similarity 33.4%; Pred. No. 4e-41;
Matches 133; Conservative 88; Mismatches 106; Indels 71; Gaps 8;
QY 3 LTVKTLKGTHTFEIRVQPNNDTMAVKNIIEEIQKDSYPWGQQLIFNGKVLKDESTLEEN 62
Db 5 I7LKTQQQTFAIRMEPDETIVKLEKIEAEKGRDAFPVAGQKLIYAGKILSDDDVPIDY 64
QY 63 KYNEDGFLVWMLSKGTSSTSSQHSNTPATRQAPLEAPQAPQPPVPAPIT----- 117
Db 65 RIDENKFNVMVT--KTRAGQGT-----APPEASPTAPESSTPPAPTSGMSHP 114
QY 118 -----TSQPEGLPAQAPNT-----HDNAASLLSGRNVDTIINLMEMG 156
Db 115 PPAAREDKSPSEASPTTSPESVSGVSSGSGREDAASTLVGTSEYETMLTEIMSG 174
QY 157 GSGWKKDKVORALRAAYNNPERAVEYLISGIPVTABIAVPIGGQANTTDRAPTGEAGLS 216
Db 175 ---YERERVVAALRASYNPNPRAVYLLTGIP-----GSPPEHSGVQESQVS 219
QY 217 GIPNTAPLDLFPQASNAGGAGGGLDLRNNPQFQAVREVMHTNPQILQPMVLVELSKQ 276
Db 220 EQPATE-----AAGENPLEFLRQPPQFQNNRQVQQNPALLPALLQQLGOE 265
QY 277 NQOILRLIENHDEFLLNNEP-----FEGEGDFLDQPEEDEMPHAISVTPEEQBAIG 330
Db 266 NPOLLQOISRHOEQFIQMLNEPPGELADISDVEGEVGAIGEAPQWNYIOVTPQEKEAIE 325
QY 331 RLESMDGDFRARIETAFACDRNEELAAAYLLEHAGEED 368
Db 326 RLKALGFPESLSVIOAYFACEKNLANFLLSQNFDE 363

RESULT 4
US-10-102-806-549
; Sequence 549, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P4103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270


```

; FILE REFERENCE: 4115-175
; CURRENT APPLICATION NUMBER: US/10/293,000
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: US 60/338,549
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 595
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-10-293-000-2

Query Match 9.2%; Score 176; DB 12; Length 595;
Best Local Similarity 21.2%; Pred. No. 7.4e-06;
Matches 99; Conservative 56; Mismatches 143; Indels 168; Gaps 19;

QY 1 MKLTVKLTAGTTFEIRVQNDTIMAVKKNIEIQGKDSYPWGQQLLIENGKVLKDESTLE 60
DB 5 IITKLTQQOTFKIRMEPEDETIVKLEKIEAKGDAFPVAGQKLIYAGKILSDDDVPIRDY 64
QY 63 KVNEDGFLVWMLSKGK 78
DB 65 RIDEKNFVVWVTKTK 80

RESULT 9
US-10-296-770-2
; Sequence 2, Application US/10296770
; Publication No. US20030104570A1
; GENERAL INFORMATION:
; APPLICANT: Cabazon Silva, Teresa Elisa Virginia
; APPLICANT: Delisse, Anne-Marie Eva Fernande
; TITLE OF INVENTION: Triple Fusion Proteins Comprising
; TITLE OF INVENTION: Ubiquitin Fused Between Thioresoxin and a Polypeptide
; TITLE OF INVENTION: Interest
; FILE REFERENCE: B45221
; CURRENT APPLICATION NUMBER: US/10/296,770
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: PCT/EP01/06952
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: GB 0015619.0
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: GB 0026484.6
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Chimaeric (B. coli - human)
US-10-296-770-2

Query Match 8.1%; Score 154; DB 15; Length 522;
Best Local Similarity 25.7%; Pred. No. 0.00039;
Matches 69; Conservative 28; Mismatches 87; Indels 84; Gaps 9;

QY 1 MKLTVKLTAGTTFEIRVQNDTIMAVKKNIEIQGKDSYPWGQQLLIENGKVLKDESTLE 60
DB 126 MQIFVKTITGKTTITVEPSDITENVK---AKIQKEGIPDQQLIPAGQLEDRGLS 182
QY 61 ENKVNEDG--FLVWMLSKGKTSGSTGTSSQHSNTPATROAPPLEAQOAPPPVAPITT 118
DB 183 DYNIOKESTLHLVRLRGW-----EQRSQH-----CKPBEGLARGEA----- 221
QY 119 SQPEGL-PAQAPNTHDNAASNLGRNVDITINQIMEMGGGSDKDKVQALRAAYNNPE 177
DB 222 ---LGLVGAQAPATEQEAASSSS---TLVE----- 246
QY 178 RAVEYLYSGIPVTAETAVPIGGGANTTDRAFTGAGLSGIPNTAPDLFPQGSNAGGG 237
DB 247 -----VTLGEVPAESPDPPSQSGASSLPTTMYPLWMSYEDSSNQ 289
QY 238 AGGGPLDFLRNNPQFOA-----VREMVH 260
DB 290 EEEGPSTFPDLESEFOALSRKVAELVH 317

RESULT 10
US-10-146-473-47
; Sequence 47, Application US/10146473
; Publication No. US2003010888A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseung

```



```

; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 47
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-473-47

Query Match
Best Local Similarity 21.0%; Score 150; DB 15; Length 624;
Matches 94; Conservative 61; Mismatches 157; Indels 136; Gaps 18;

Qy 1 MKLTVKTLKGTHTPEIRVQNDTIMAVKKNIEIQGKDSYPMGQQLLIENGKVLKDESTLE 60
Db 33 IKVTTPKEKE-EFVAPENSSVQPKF---EALSKEFKSQTDQLVLIPAGKILKQDITLI 88
Qy 61 ENKVNEDGFLVMSLKGK-----TSGST----- 83
Db 89 QHGIIH-DGLTVHLVIKSNRPQGSTQPSNAAGTNTTSASTPRSNSTPISTNSNPFGLGS 147
Qy 84 -----GTSSQHSNTPATROAPPLEAQ---QAPQPPVAPITTSOPEGL----- 124
Db 148 LGLAGLSSLGLSTNFSLSQSQMQLWASPEMMIQIMENFVQSMLSNPLMRQLIMA 207
Qy 125 --PAQANTHDAASLLSG-----RNVDTIINQLMEMGGGSWD 161
Db 208 NPQMQLQIQRNPEISHLLNPDIMRQTLRIARNPAMQWEMENQDLALSNSLESIFGG--- 264
Qy 162 KDKVQALRAAYNPE-----RAVEYLYSGIPYTAIEIPIVGGQ---ANTTDRAPTGEAG 214
Db 265 ---YNALRRMYTDIQEPLNAAQEQFGNPFASVSGSSSSGEGTQPSRTENRDP----- 315
Qy 215 LSGIPNT-APLDLDFPGA-----SNAGGCGAGGGLDPLFN-----NPQFOAV 255
Db 316 ---LPNPWAPPPATQSSATSTTTTSGSGSGSSNATGNTVAANYVASIPSTPGMOSL 372
Qy 256 REMVHTNPQILQPL-----VELSKQNPQILRLIENHDEFL--QLLANEPFEGSGD 305
Db 373 LQQTENPQILQNLSPAYMRSNMWQSLSQNPDLAQMMLNSPLFTANPQLQEQMRPLPA 432
Qy 306 FLQPEDEMPHAISVTPPEQBAIGLE 333
Db 433 FLQMQNPDTLSAMS-NPRAMQALMQIQ 459

RESULT 11
US-09-938-864-410
; Sequence 410, Application US/09938864
; Publication No. US20030072767A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Smithgall, Molly
; APPLICANT: Moulton, Gus
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Mossman, Sally
; APPLICANT: Evans, Lawrence
; APPLICANT: Spies, A. Gregory
; APPLICANT: Boydston, Jeremy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
; FILE REFERENCE: 210121.465C5
; CURRENT APPLICATION NUMBER: US/09/938,864
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 413

; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 410
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-195-835-410

Query Match
Best Local Similarity 30.0%; Score 140; DB 12; Length 504;
Matches 48; Conservative 15; Mismatches 59; Indels 38; Gaps 4

Qy 1 MKLTVKTLKGTHTPEIRVQNDTIMAVKKNIEIQGKDSYPMGQQLLIENGKVLKDESTLE 60
Db 1 MQIFVKTLTGKTTILEVPSDTIENVK---AKIQKEGIPPDQORLIPAGKQLEDRGRTLS 57
Qy 61 ENKVNEDGFLVVML-----SKGTSSTGTSTSSQHSNTPATROA 99
Db 58 DYNIQKESTLHLVRLRGAMGSDVRDLNALLPAVPSLGGGGCALPVSQAQWAPVLDF 117
Qy 100 PP-----LEAPQAPQPPVAPITTSOPEGLPAQAPN 130
Db 118 PPGASAYGSLGGPAPPAPPPPPPP-----PHSFIKQEPS 152

RESULT 12
US-10-195-835-410
; Sequence 410, Application US/10195835
; Publication No. US20030198622A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Smithgall, Molly D.
; APPLICANT: Carter, Darrick
; APPLICANT: Cheever, Martin A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Sutherland, R. Alec
; APPLICANT: Mossman, Sally P.
; APPLICANT: Evans, Lawrence S.
; APPLICANT: Swanson, Ryan M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
; FILE REFERENCE: 210121.465C8
; CURRENT APPLICATION NUMBER: US/10/195,835
; CURRENT FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 410
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-195-835-410

Query Match
Best Local Similarity 30.0%; Score 140; DB 12; Length 504;
Matches 48; Conservative 15; Mismatches 59; Indels 38; Gaps 4

Qy 1 MKLTVKTLKGTHTPEIRVQNDTIMAVKKNIEIQGKDSYPMGQQLLIENGKVLKDESTLE 60
Db 1 MQIFVKTLTGKTTILEVPSDTIENVK---AKIQKEGIPPDQORLIPAGKQLEDRGRTLS 57
Qy 61 ENKVNEDGFLVVML-----SKGTSSTGTSTSSQHSNTPATROA 99
Db 58 DYNIQKESTLHLVRLRGAMGSDVRDLNALLPAVPSLGGGGCALPVSQAQWAPVLDF 117
Qy 100 PP-----LEAPQAPQPPVAPITTSOPEGLPAQAPN 130
Db 118 PPGASAYGSLGGPAPPAPPPPPPP-----PHSFIKQEPS 152

RESULT 13
US-10-286-333-410
; Sequence 410, Application US/10286333
; Publication No. US20030215458A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Gaiger, Alexander
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Jaya, No. US20030215458Alalalie
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
; TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
; FILE REFERENCE: 210121.465C10
; CURRENT APPLICATION NUMBER: US/10/286,333
; CURRENT FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 410
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-286-333-410

Query Match 7.3%; Score 140; DB 12; Length 504;
Best Local Similarity 30.0%; Pred. No. 0.0052;
Matches 48; Conservative 15; Mismatches 59; Indels 38; Gaps 4;

Qy 1 MKLTVKTLKGFHEIRVQPNDDTMAVKNIIEIQKDSYPMWGOQLLIENGKVLKDESTLE 60
Db 1 MQIFVKLTGKTTITLEVEPSDTIENVK---AKIQDKGIPPDQORLIFAGKQLEDGRTL 57
Qy 61 ENKVNEDGFLVVML-----SKGKTSSTGTSSSQHSNTPATRQA 99
Db 58 DYNIQESTLHLVLRGAMGSDVRDLNALLPAVPSLGGGGGCGALPVSGAAQWAPVLDF 117
Qy 100 PP-----LEAPQAPQPPVAPITTSQPEGLPAQAPN 130
Db 118 PPGASAYGSLGGPAPPAPPPPPPP-----PHSFIKQEPS 152

RESULT 14
US-10-125-635A-410
; Sequence 410, Application US/10125635A
; Publication No. US20030039635A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Smithgall, Molly D.
; APPLICANT: Carter, Darrick
; APPLICANT: Cheever, Martin A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Sutherland, R. Alec
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
; TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
; FILE REFERENCE: 210121.465C7
; CURRENT APPLICATION NUMBER: US/10/125,635A
; CURRENT FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 410
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-125-635A-410

Query Match 7.3%; Score 140; DB 15; Length 504;
Best Local Similarity 30.0%; Pred. No. 0.0052;
Matches 48; Conservative 15; Mismatches 59; Indels 38; Gaps 4;

Qy 1 MKLTVKTLKGFHEIRVQPNDDTMAVKNIIEIQKDSYPMWGOQLLIENGKVLKDESTLE 60
Db 1 MQIFVKLTGKTTITLEVEPSDTIENVK---AKIQDKGIPPDQORLIFAGKQLEDGRTL 57
Qy 61 ENKVNEDGFLVVML-----SKGKTSSTGTSSSQHSNTPATRQA 99
Db 58 DYNIQESTLHLVLRGAMGSDVRDLNALLPAVPSLGGGGGCGALPVSGAAQWAPVLDF 117
Qy 100 PP-----LEAPQAPQPPVAPITTSQPEGLPAQAPN 130
Db 118 PPGASAYGSLGGPAPPAPPPPPPP-----PHSFIKQEPS 152

RESULT 15
US-10-002-603-410
; Sequence 410, Application US/10002603
; Publication No. US20030095971A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Smithgall, Molly
; APPLICANT: Moulton, Gus
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Mossman, Sally
; APPLICANT: Evans, Lawrence
; APPLICANT: Spies, A. Gregory
; APPLICANT: Boydston, Jeremy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
; TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
; FILE REFERENCE: 210121.465C6
; CURRENT APPLICATION NUMBER: US/10/002,603
; CURRENT FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 413
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 410
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-002-603-410

Query Match 7.3%; Score 140; DB 15; Length 504;
Best Local Similarity 30.0%; Pred. No. 0.0052;
Matches 48; Conservative 15; Mismatches 59; Indels 38; Gaps 4;

Qy 1 MKLTVKTLKGFHEIRVQPNDDTMAVKNIIEIQKDSYPMWGOQLLIENGKVLKDESTLE 60
Db 1 MQIFVKLTGKTTITLEVEPSDTIENVK---AKIQDKGIPPDQORLIFAGKQLEDGRTL 57
Qy 61 ENKVNEDGFLVVML-----SKGKTSSTGTSSSQHSNTPATRQA 99
Db 58 DYNIQESTLHLVLRGAMGSDVRDLNALLPAVPSLGGGGGCGALPVSGAAQWAPVLDF 117
Qy 100 PP-----LEAPQAPQPPVAPITTSQPEGLPAQAPN 130
Db 118 PPGASAYGSLGGPAPPAPPPPPPP-----PHSFIKQEPS 152

Search completed: December 17, 2003, 06:31:56
Job time : 127.158 secs

```

```

; APPLICANT: Gaiger, Alexander
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Jaya, No. US20030215458Alalalie
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
; TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
; FILE REFERENCE: 210121.465C10
; CURRENT APPLICATION NUMBER: US/10/286,333
; CURRENT FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 410
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-286-333-410

Query Match 7.3%; Score 140; DB 12; Length 504;
Best Local Similarity 30.0%; Pred. No. 0.0052;
Matches 48; Conservative 15; Mismatches 59; Indels 38; Gaps 4;

Qy 1 MKLTVKTLKGFHEIRVQPNDDTMAVKNIIEIQKDSYPMWGOQLLIENGKVLKDESTLE 60
Db 1 MQIFVKLTGKTTITLEVEPSDTIENVK---AKIQDKGIPPDQORLIFAGKQLEDGRTL 57
Qy 61 ENKVNEDGFLVVML-----SKGKTSGSTGTSSSQHSNTPATRQA 99
Db 58 DYNIQESTLHLVLRGAMGSDVRDLNALLPAVPSLGGGGGCGALPVSGAAQWAPVLDF 117
Qy 100 PP-----LEAPQAPQPPVAPITTSQPEGLPAQAPN 130
Db 118 PPGASAYGSLGGPAPPAPPPPPPP-----PHSFIKQEPS 152

RESULT 14
US-10-125-635A-410
; Sequence 410, Application US/10125635A
; Publication No. US20030039635A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Smithgall, Molly D.
; APPLICANT: Carter, Darrick
; APPLICANT: Cheever, Martin A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Sutherland, R. Alec
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
; TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
; FILE REFERENCE: 210121.465C7
; CURRENT APPLICATION NUMBER: US/10/125,635A
; CURRENT FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 410
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-125-635A-410

Query Match 7.3%; Score 140; DB 15; Length 504;
Best Local Similarity 30.0%; Pred. No. 0.0052;
Matches 48; Conservative 15; Mismatches 59; Indels 38; Gaps 4;

Qy 1 MKLTVKTLKGFHEIRVQPNDDTMAVKNIIEIQKDSYPMWGOQLLIENGKVLKDESTLE 60
Db 1 MQIFVKLTGKTTITLEVEPSDTIENVK---AKIQDKGIPPDQORLIFAGKQLEDGRTL 57
Qy 61 ENKVNEDGFLVVML-----SKGKTSGSTGTSSSQHSNTPATRQA 99
Db 58 DYNIQESTLHLVLRGAMGSDVRDLNALLPAVPSLGGGGGCGALPVSGAAQWAPVLDF 117
Qy 100 PP-----LEAPQAPQPPVAPITTSQPEGLPAQAPN 130
Db 118 PPGASAYGSLGGPAPPAPPPPPPP-----PHSFIKQEPS 152

RESULT 15
US-10-002-603-410
; Sequence 410, Application US/10002603
; Publication No. US20030095971A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Smithgall, Molly
; APPLICANT: Moulton, Gus
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Mossman, Sally
; APPLICANT: Evans, Lawrence
; APPLICANT: Spies, A. Gregory
; APPLICANT: Boydston, Jeremy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
; TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
; FILE REFERENCE: 210121.465C6
; CURRENT APPLICATION NUMBER: US/10/002,603
; CURRENT FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 413
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 410
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-002-603-410

Query Match 7.3%; Score 140; DB 15; Length 504;
Best Local Similarity 30.0%; Pred. No. 0.0052;
Matches 48; Conservative 15; Mismatches 59; Indels 38; Gaps 4;

Qy 1 MKLTVKTLKGFHEIRVQPNDDTMAVKNIIEIQKDSYPMWGOQLLIENGKVLKDESTLE 60
Db 1 MQIFVKLTGKTTITLEVEPSDTIENVK---AKIQDKGIPPDQORLIFAGKQLEDGRTL 57
Qy 61 ENKVNEDGFLVVML-----SKGKTSGSTGTSSSQHSNTPATRQA 99
Db 58 DYNIQESTLHLVLRGAMGSDVRDLNALLPAVPSLGGGGGCGALPVSGAAQWAPVLDF 117
Qy 100 PP-----LEAPQAPQPPVAPITTSQPEGLPAQAPN 130
Db 118 PPGASAYGSLGGPAPPAPPPPPPP-----PHSFIKQEPS 152

Search completed: December 17, 2003, 06:31:56
Job time : 127.158 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 06:17:26 ; Search time 129.014 Seconds
(without alignments)
2595.451 Million cell updates/sec

Title: US-09-805-550-4

Perfect score: 1910

Sequence: 1 MKLTVTLKGFHFIRVQPN.....CDRNEELAANYLLEHAGEED 368

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main.*

1: /cgn2_6/ptodata/1/paa/PCTUS COMB.pcp.*
2: /cgn2_6/ptodata/1/paa/US06 COMB.pcp.*
3: /cgn2_6/ptodata/1/paa/US07 COMB.pcp.*
4: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
5: /cgn2_6/ptodata/1/paa/US081 COMB.pcp.*
6: /cgn2_6/ptodata/1/paa/US082 COMB.pcp.*
7: /cgn2_6/ptodata/1/paa/US083 COMB.pcp.*
8: /cgn2_6/ptodata/1/paa/US084 COMB.pcp.*
9: /cgn2_6/ptodata/1/paa/US085 COMB.pcp.*
10: /cgn2_6/ptodata/1/paa/US086 COMB.pcp.*
11: /cgn2_6/ptodata/1/paa/US087 COMB.pcp.*
12: /cgn2_6/ptodata/1/paa/US088 COMB.pcp.*
13: /cgn2_6/ptodata/1/paa/US089 COMB.pcp.*
14: /cgn2_6/ptodata/1/paa/US090 COMB.pcp.*
15: /cgn2_6/ptodata/1/paa/US091 COMB.pcp.*
16: /cgn2_6/ptodata/1/paa/US092 COMB.pcp.*
17: /cgn2_6/ptodata/1/paa/US093 COMB.pcp.*
18: /cgn2_6/ptodata/1/paa/US094 COMB.pcp.*
19: /cgn2_6/ptodata/1/paa/US095 COMB.pcp.*
20: /cgn2_6/ptodata/1/paa/US096 COMB.pcp.*
21: /cgn2_6/ptodata/1/paa/US097A COMB.pcp.*
22: /cgn2_6/ptodata/1/paa/US097B COMB.pcp.*
23: /cgn2_6/ptodata/1/paa/US098 COMB.pcp.*
24: /cgn2_6/ptodata/1/paa/US099A COMB.pcp.*
25: /cgn2_6/ptodata/1/paa/US099B COMB.pcp.*
26: /cgn2_6/ptodata/1/paa/US100 COMB.pcp.*
27: /cgn2_6/ptodata/1/paa/US101 COMB.pcp.*
28: /cgn2_6/ptodata/1/paa/US102 COMB.pcp.*
29: /cgn2_6/ptodata/1/paa/US103 COMB.pcp.*
30: /cgn2_6/ptodata/1/paa/US104 COMB.pcp.*
31: /cgn2_6/ptodata/1/paa/US106 COMB.pcp.*
32: /cgn2_6/ptodata/1/paa/US60 COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1910	100.0	368	21	US-09-708-427-51368 Sequence 51368, A

2	1910	100.0	368	23	US-09-805-550-4	Sequence 4, Appli
3	1910	100.0	368	32	US-60-109-728-4	Sequence 4, Appli
4	1792	93.8	246	21	US-09-708-427-51369	Sequence 51369, A
5	1540	80.6	296	21	US-09-708-427-51370	Sequence 51370, A
6	1524.5	79.8	369	30	US-10-437-963-161952	Sequence 161952, A
7	1288.5	67.5	252	28	US-10-219-999-49896	Sequence 49896, A
8	1288.5	67.5	289	30	US-10-425-114-59208	Sequence 59208, A
9	1185	62.0	363	30	US-10-424-599-170467	Sequence 170467, A
10	1169.5	61.2	371	19	US-09-513-996A-18372	Sequence 18372, A
11	1169.5	61.2	363	30	US-09-620-394B-6541	Sequence 6541, Ap
12	1168.5	61.2	371	20	US-10-424-599-219691	Sequence 219691, A
13	1133.5	60.4	379	22	US-09-791-537-55742	Sequence 55742, A
14	1115	58.4	368	19	US-09-513-996A-55531	Sequence 55531, A
15	1067.5	55.9	348	19	US-09-513-996A-18373	Sequence 18373, A
16	1067.5	55.9	348	20	US-09-620-394B-6542	Sequence 6542, Ap
17	1043.5	54.6	367	22	US-09-791-537-18661	Sequence 18661, A
18	1019	53.4	389	22	US-09-791-537-58054	Sequence 58054, A
19	1014	53.1	345	19	US-09-513-996A-55532	Sequence 55532, A
20	1013.5	53.1	400	28	US-10-219-999-35327	Sequence 35327, A
21	1013.5	53.1	400	30	US-10-424-599-256715	Sequence 256715, A
22	1013.5	53.1	400	32	US-60-324-109-20739	Sequence 20739, A
23	1013.5	53.1	433	30	US-10-425-114-45576	Sequence 45576, A
24	1006.5	52.7	403	30	US-10-424-599-200504	Sequence 200504, A
25	1004.5	52.6	392	28	US-10-219-999-43947	Sequence 43947, A
26	1004.5	52.6	425	30	US-10-425-114-54716	Sequence 54716, A
27	1003.5	52.5	392	30	US-10-424-599-256719	Sequence 256719, A
28	978	51.2	419	22	US-09-791-537-66481	Sequence 66481, A
29	978	51.2	419	24	US-09-935-625-8508	Sequence 8508, Ap
30	954	49.9	382	22	US-09-791-537-55721	Sequence 55721, A
31	931.5	48.8	378	19	US-09-513-996A-21981	Sequence 21981, A
32	931.5	48.8	378	19	US-09-513-996A-56723	Sequence 56723, A
33	912	47.7	402	28	US-10-219-999-62455	Sequence 62455, A
34	912	47.7	402	32	US-60-324-109-28561	Sequence 28561, A
35	912	47.7	402	32	US-60-324-109-28562	Sequence 28562, A
36	912	47.7	424	30	US-10-425-114-72275	Sequence 72275, A
37	908.5	47.6	405	28	US-10-219-999-39045	Sequence 39045, A
38	908.5	47.6	405	28	US-10-219-999-50088	Sequence 50088, A
39	908.5	47.6	405	32	US-60-312-544-8573	Sequence 8573, Ap
40	908.5	47.6	418	30	US-10-425-114-41386	Sequence 41386, A
41	908.5	47.6	418	30	US-10-425-114-58736	Sequence 58736, A
42	904	47.3	442	28	US-10-219-999-55396	Sequence 55396, A
43	904	47.3	442	30	US-10-425-114-64677	Sequence 64677, A
44	903	47.3	402	28	US-10-219-999-54143	Sequence 54143, A
45	903	47.3	424	30	US-10-425-114-59667	Sequence 59667, A

ALIGNMENTS

RESULT 1

US-09-708-427-51368
; Sequence 51368, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708, 427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51368
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..368
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: 1..368
; OTHER INFORMATION: Ceres Seq. ID 1926693
US-09-708-427-51368

```
Query Match      100.0%; Score 1910; DB 21; Length 368;
Best Local Similarity 100.0%; Pred. No. 9.2e-149;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKLTVKTLKGTHFEIRVQPNNDTMAVKKNIEIQKDSYPWGQQLLIENGKVLKDESTLE 60
Db 1 MKLTVKTLKGTHFEIRVQPNNDTMAVKKNIEIQKDSYPWGQQLLIENGKVLKDESTLE 60

Qy 61 ENKVNEDGLVVLMSKGTSGTSSQHSNTPATROAPPLEAQPPVAPITTSQ 120
Db 61 ENKVNEDGLVVLMSKGTSGTSSQHSNTPATROAPPLEAQPPVAPITTSQ 120

Qy 121 PEGLPQAQPNTHDAAASLLSGRNVDTTIINQLEMGGSGWDKQVQRALRAAYNNPERAV 180
Db 121 PEGLPQAQPNTHDAAASLLSGRNVDTTIINQLEMGGSGWDKQVQRALRAAYNNPERAV 180

Qy 181 EYLYSGIPVTAEIAPVIGGGANTTDRAPTGEAGLSGIPNTAPLDLFPQGASNAGGAGG 240
Db 181 EYLYSGIPVTAEIAPVIGGGANTTDRAPTGEAGLSGIPNTAPLDLFPQGASNAGGAGG 240

Qy 241 GPLDFLRNNPQOAVREMVHTNPQIILQPMVLVELSKONPQILRIEENHDEFLLQNEPFE 300
Db 241 GPLDFLRNNPQOAVREMVHTNPQIILQPMVLVELSKONPQILRIEENHDEFLLQNEPFE 300

Qy 301 GCGEDFLDQPEDEMHPHAI SVTPEEQEAI GRLESMDGDRARVIEAF LACDRNEELAANYL 360
Db 301 GCGEDFLDQPEDEMHPHAI SVTPEEQEAI GRLESMDGDRARVIEAF LACDRNEELAANYL 360

Qy 361 LEHAGED 368
Db 361 LEHAGED 368

RESULT 2
US-09-805-550-4
; Sequence 4, Application US/0980550
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; APPLICANT: Tagliani, Laura
; APPLICANT: Bowen, Benjamin
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Rad23 Genes and Uses Thereof
; FILE REFERENCE: 0964P
; CURRENT APPLICATION NUMBER: US/09/805,550
; CURRENT FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 09/413,574
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: 60/109,728
; PRIOR FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Zea mays
US-09-805-550-4

Query Match      100.0%; Score 1910; DB 23; Length 368;
Best Local Similarity 100.0%; Pred. No. 9.2e-149;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKLTVKTLKGTHFEIRVQPNNDTMAVKKNIEIQKDSYPWGQQLLIENGKVLKDESTLE 60
Db 1 MKLTVKTLKGTHFEIRVQPNNDTMAVKKNIEIQKDSYPWGQQLLIENGKVLKDESTLE 60

Qy 61 ENKVNEDGLVVLMSKGTSGTSSQHSNTPATROAPPLEAQPPVAPITTSQ 120
Db 61 ENKVNEDGLVVLMSKGTSGTSSQHSNTPATROAPPLEAQPPVAPITTSQ 120

Qy 121 PEGLPQAQPNTHDAAASLLSGRNVDTTIINQLEMGGSGWDKQVQRALRAAYNNPERAV 180
Db 121 PEGLPQAQPNTHDAAASLLSGRNVDTTIINQLEMGGSGWDKQVQRALRAAYNNPERAV 180

Qy 181 EYLYSGIPVTAEIAPVIGGGANTTDRAPTGEAGLSGIPNTAPLDLFPQGASNAGGAGG 240
Db 181 EYLYSGIPVTAEIAPVIGGGANTTDRAPTGEAGLSGIPNTAPLDLFPQGASNAGGAGG 240

Query Match      100.0%; Score 1910; DB 32; Length 368;
Best Local Similarity 100.0%; Pred. No. 9.2e-149;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKLTVKTLKGTHFEIRVQPNNDTMAVKKNIEIQKDSYPWGQQLLIENGKVLKDESTLE 60
Db 1 MKLTVKTLKGTHFEIRVQPNNDTMAVKKNIEIQKDSYPWGQQLLIENGKVLKDESTLE 60

Qy 61 ENKVNEDGLVVLMSKGTSGTSGTSSQHSNTPATROAPPLEAQPPVAPITTSQ 120
Db 61 ENKVNEDGLVVLMSKGTSGTSGTSSQHSNTPATROAPPLEAQPPVAPITTSQ 120

Qy 121 PEGLPQAQPNTHDAAASLLSGRNVDTTIINQLEMGGSGWDKQVQRALRAAYNNPERAV 180
Db 121 PEGLPQAQPNTHDAAASLLSGRNVDTTIINQLEMGGSGWDKQVQRALRAAYNNPERAV 180

Qy 181 EYLYSGIPVTAEIAPVIGGGANTTDRAPTGEAGLSGIPNTAPLDLFPQGASNAGGAGG 240
Db 181 EYLYSGIPVTAEIAPVIGGGANTTDRAPTGEAGLSGIPNTAPLDLFPQGASNAGGAGG 240

Qy 241 GPLDFLRNNPQOAVREMVHTNPQIILQPMVLVELSKONPQILRIEENHDEFLLQNEPFE 300
Db 241 GPLDFLRNNPQOAVREMVHTNPQIILQPMVLVELSKONPQILRIEENHDEFLLQNEPFE 300

Qy 301 GCGEDFLDQPEDEMHPHAI SVTPEEQEAI GRLESMDGDRARVIEAF LACDRNEELAANYL 360
Db 301 GCGEDFLDQPEDEMHPHAI SVTPEEQEAI GRLESMDGDRARVIEAF LACDRNEELAANYL 360

Qy 361 LEHAGED 368
Db 361 LEHAGED 368

RESULT 3
US-60-109-728-4
; Sequence 4, Application US/60109728
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; APPLICANT: Tagliani, Laura
; APPLICANT: Bowen, Benjamin
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Rad23 Genes and Uses Thereof
; FILE REFERENCE: 0964P
; CURRENT APPLICATION NUMBER: US/60/109,728
; CURRENT FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Zea mays
US-60-109-728-4

Query Match      100.0%; Score 1910; DB 32; Length 368;
Best Local Similarity 100.0%; Pred. No. 9.2e-149;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKLTVKTLKGTHFEIRVQPNNDTMAVKKNIEIQKDSYPWGQQLLIENGKVLKDESTLE 60
Db 1 MKLTVKTLKGTHFEIRVQPNNDTMAVKKNIEIQKDSYPWGQQLLIENGKVLKDESTLE 60

Qy 61 ENKVNEDGLVVLMSKGTSGTSGTSSQHSNTPATROAPPLEAQPPVAPITTSQ 120
Db 61 ENKVNEDGLVVLMSKGTSGTSGTSSQHSNTPATROAPPLEAQPPVAPITTSQ 120

Qy 121 PEGLPQAQPNTHDAAASLLSGRNVDTTIINQLEMGGSGWDKQVQRALRAAYNNPERAV 180
Db 121 PEGLPQAQPNTHDAAASLLSGRNVDTTIINQLEMGGSGWDKQVQRALRAAYNNPERAV 180

Qy 181 EYLYSGIPVTAEIAPVIGGGANTTDRAPTGEAGLSGIPNTAPLDLFPQGASNAGGAGG 240
Db 181 EYLYSGIPVTAEIAPVIGGGANTTDRAPTGEAGLSGIPNTAPLDLFPQGASNAGGAGG 240

Qy 241 GPLDFLRNNPQOAVREMVHTNPQIILQPMVLVELSKONPQILRIEENHDEFLLQNEPFE 300
Db 241 GPLDFLRNNPQOAVREMVHTNPQIILQPMVLVELSKONPQILRIEENHDEFLLQNEPFE 300

Qy 301 GCGEDFLDQPEDEMHPHAI SVTPEEQEAI GRLESMDGDRARVIEAF LACDRNEELAANYL 360
Db 301 GCGEDFLDQPEDEMHPHAI SVTPEEQEAI GRLESMDGDRARVIEAF LACDRNEELAANYL 360

Qy 361 LEHAGED 368
Db 361 LEHAGED 368

RESULT 4
US-09-708-427-51369
; Sequence 427-51369, Application US/09708427
; GENERAL INFORMATION:
```

```

; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51369
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..345
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..345
; OTHER INFORMATION: Ceres Seq. ID 1926694
; US-09-708-427-51369

Query Match      93.8%; Score 1792; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 4.8e-139;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 MAVKNIIEIOGKSYPMGQQLIFNGKVLKDESTLEENKVNEDGFLVVLMSKGTSGST 83
Db 1 MAVKNIIEIOGKSYPMGQQLIFNGKVLKDESTLEENKVNEDGFLVVLMSKGTSGST 60

Qy 84 GTSSSOHSNTATRAQAPPLAPQAPPPVAPITTSQPEGLPAQAPNTHDNAANLLSGR 143
Db 61 GTSSSOHSNTATRAQAPPLAPQAPPPVAPITTSQPEGLPAQAPNTHDNAANLLSGR 120

Qy 144 NVDTIINQLMEMGGGSKDKVQRALRAAYNNPERAVEYLISGIPVTAIEIAPVIGGGGAN 203
Db 121 NVDTIINQLMEMGGGSKDKVQRALRAAYNNPERAVEYLISGIPVTAIEIAPVIGGGGAN 180

Qy 204 TTDRAPTGEAGLSGIPNTAPLDLPFGQASNAGGAGGGGGLFLRNNPQFQAVREMVHTNP 263
Db 181 TTDRAPTGEAGLSGIPNTAPLDLPFGQASNAGGAGGGGGLFLRNNPQFQAVREMVHTNP 240

Qy 264 QILQPMVLVSKNPQILRLIEENHDEFLQLNPPFEGGEGDFLDQPEDEMPHAISVTP 323
Db 241 QILQPMVLVSKNPQILRLIEENHDEFLQLNPPFEGGEGDFLDQPEDEMPHAISVTP 300

Qy 324 EEQEAIGRLSMGDFRARIIEAFACDRNEELAANYLLEHAGEED 368
Db 301 EEQEAIGRLSMGDFRARIIEAFACDRNEELAANYLLEHAGEED 345

RESULT 5
US-09-708-427-51370
; Sequence 51370, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51370
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..296
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..296
; OTHER INFORMATION: Ceres Seq. ID 1926695
; US-09-708-427-51370

Query Match      80.6%; Score 1540; DB 21; Length 296;
Best Local Similarity 100.0%; Pred. No. 2.6e-118;
Matches 296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 73 MLSKGKTSSTGTSSQHSNTPATRAQAPPLAPQAPPPVAPITTSQPEGLPAQAPNTH 132
Db 1 MLSKGKTSSTGTSSQHSNTPATRAQAPPLAPQAPPPVAPITTSQPEGLPAQAPNTH 60

Qy 133 DNAASNLLSGRNVDTIINQLMEMGGGSKDKVQRALRAAYNNPERAVEYLISGIPVTAIE 192
Db 61 DNAASNLLSGRNVDTIINQLMEMGGGSKDKVQRALRAAYNNPERAVEYLISGIPVTAIE 120

Qy 193 IAVPIGGQAGNTTDRAPTGEAGLSGIPNTAPLDLPFGQASNAGGAGGGGGLFLRNNPQF 252
Db 121 IAVPIGGQAGNTTDRAPTGEAGLSGIPNTAPLDLPFGQASNAGGAGGGGGLFLRNNPQF 180

Qy 253 QAVREMVHTNPQILQPMVLVSKNPQILRLIEENHDEFLQLNPPFEGGEGDFLDQPEE 312
Db 181 QAVREMVHTNPQILQPMVLVSKNPQILRLIEENHDEFLQLNPPFEGGEGDFLDQPEE 240

Qy 313 DEMPHAISVTPPEQEAIGRLSMGDFRARIIEAFACDRNEELAANYLLEHAGEED 368
Db 241 DEMPHAISVTPPEQEAIGRLSMGDFRARIIEAFACDRNEELAANYLLEHAGEED 296

RESULT 6
US-10-437-963-161952
; Sequence 161952, Application US/10437963
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yibua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 161952
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_6108C.1.pep
; US-10-437-963-161952

Query Match      79.8%; Score 1524.5; DB 30; Length 369;
Best Local Similarity 79.9%; Pred. No. 7e-117;
Matches 295; Conservative 31; Mismatches 42; Indels 1; Gaps 1;

Qy 1 MKLTVKTLKGTHTFEIRVQPNDDTMAVKNIIEIOGKSYPMGQQLIFNGKVLKDESTLE 60
Db 1 MKLTVKTLKGTHTFEIRVQPNDDTMAVKNIIEIOGKSYPMGQQLIFNGKVLKDESTLE 60

Qy 61 ENKVNEDGFLVVLMSKGTSGTSSQHSNTPATRAQAPPLAPQAPPPVAPITTSQ 120
Db 61 ENKVNEDGFLVVLMSKGTSGTSSQHSNTPATRAQAPPLAPQAPPPVAPITTSQ 120

Qy 121 PEGLPAAQAP-NTHDNAASNLLSGRNVDTIINQLMEMGGGSKDKVQRALRAAYNNPERA 179
Db 121 PERPPAEAPSNAVQQAASNLLSGNLDTTINQLMEMGGGSKDRDKVQRALRAAYNNPERA 180

Qy 180 VEYLISGIPVTAIEIAPVIGGGANTTDRAPTGEAGLSGIPNTAPLDLPFGQASNAGGAG 239
Db 181 VEYLISGIPVTAIEIAPVIGGGANTTDRAPTGEAGLSGIPNTAPLDLPFGQASNAGGAG 240

```

```
QY 240 GGPLDLFLRNPPQFQAVRMVHTNPQILOPMLVELSKNPQILRLIBENHDFLOLLNEPP 299
Db 241 GGTLEFLRHNOQFQALREVMHTNPQILOPMLQELSKNPQILRLIQENHDFLOLLNEPP 300
QY 300 EGGEDGFLDQDEEMPHASIVTPEEQEAGRLSGMGFDRARVIEAFACDRNBEAANY 359
Db 301 DGADGDFLDQDQDEMPHSINVTPEEQEAGRLSGMGFDRARVIEAFACDRNEOLAANY 360
QY 360 LLEHAGEED 368
Db 361 LLEHADED 369

RESULT 7
US-10-219-999-49896
; Sequence 49896, Application US/10219999
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)C
; CURRENT APPLICATION NUMBER: US/10/219,999
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/324,109
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/312,544
; PRIOR FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 63520
; SEQ ID NO 49896
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Zea mays
US-10-219-999-49896

Query Match 67.5%; Score 1288.5; DB 28; Length 252;
Best Local Similarity 99.2%; Pred. No. 1.3e-97;
Matches 250; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 MKLTVTKLKGTHFEIRVQPNNDTMAVKKNIEEIQGKDSYPWGQQLLIFNGKVLKDESTLE 60
Db 1 MKLTVTKLKGTHFEIRVQPNNDTMAVKKNIEEIQGKDSYPWGQQLLIFNGKVLKDESTLE 60
QY 61 ENKVNEDGFLVWMLSKGKTSSTGTSQHSNTPATROAPPLEAPQOAPPPVAPITTSQ 120
Db 61 ENKVNEDGFLVWMLSKGKTSSTGTSQHSNTPATROAPPLEAPQOAPPPVAPITTSQ 120
QY 121 PEGIPAQAPNTHDAAASNLSSGRNVDTIINQLMEMGGSGWDKQVQRALRAAYNNPERAV 180
Db 121 PEGIPAQAPNTHDAAASNLSSGRNVDTIINQLMEMGGSGWDKQVQRALRAAYNNPERAV 180
QY 181 EYLSGIPVTAEIAVPIGGQGANTTDRAPTGEAGLSGIPNTAPLDLFF-QGASNAGGAG 239
Db 181 EYLSGIPVTAEIAVPIGGQGANTTDRAPTGEAGLSGIPNTAPLDLFF-QGASNAGGAG 239
QY 240 GGPLDLFLRNPPQ 251
Db 241 GGPLDLFLRNPPQ 252

RESULT 8
US-10-425-114-59208
; Sequence 59208, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
```

```
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59208
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700264466_FLI.pep
US-10-425-114-59208

Query Match 67.5%; Score 1288.5; DB 30; Length 289;
Best Local Similarity 99.2%; Pred. No. 1.6e-97;
Matches 250; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 MKLTVTKLKGTHFEIRVQPNNDTMAVKKNIEEIQGKDSYPWGQQLLIFNGKVLKDESTLE 60
Db 30 MKLTVTKLKGTHFEIRVQPNNDTMAVKKNIEEIQGKDSYPWGQQLLIFNGKVLKDESTLE 89
QY 61 ENKVNEDGFLVWMLSKGKTSSTGTSQHSNTPATROAPPLEAPQOAPPPVAPITTSQ 120
Db 90 ENKVNEDGFLVWMLSKGKTSSTGTSQHSNTPATROAPPLEAPQOAPPPVAPITTSQ 149
QY 121 PEGIPAQAPNTHDAAASNLSSGRNVDTIINQLMEMGGSGWDKQVQRALRAAYNNPERAV 180
Db 150 PEGIPAQAPNTHDAAASNLSSGRNVDTIINQLMEMGGSGWDKQVQRALRAAYNNPERAV 209
QY 181 EYLSGIPVTAEIAVPIGGQGANTTDRAPTGEAGLSGIPNTAPLDLFF-QGASNAGGAG 239
Db 210 EYLSGIPVTAEIAVPIGGQGANTTDRAPTGEAGLSGIPNTAPLDLFF-QGASNAGGAG 269
QY 240 GGPLDLFLRNPPQ 251
Db 270 GGPLDLFLRNPPQ 281

RESULT 9
US-10-424-599-170467
; Sequence 170467, Application US/10424599
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 170467
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(363)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_124948C.1.pap
US-10-424-599-170467

Query Match 62.0%; Score 1185; DB 30; Length 363;
Best Local Similarity 64.0%; Pred. No. 8.2e-89;
Matches 236; Conservative 51; Mismatches 70; Indels 12; Gaps 5;

QY 1 MKLTVTKLKGTHFEIRVQPNNDTMAVKKNIEEIQGKDSYPWGQQLLIFNGKVLKDESTLE 60
Db 1 MKLTVTKLKGTHFEIRVQPNNDTMAVKKNIEEIQGKDSYPWGQQLLIFNGKVLKDESTLE 60
```

```
Qy 61 ENKVEDGLVWMLSGKTSSTGSSQHSNTPATRQAPLEAPQAP--QPPVAPITT 118
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 DNKVSDEGLVWMLSGKTLGAGSSTQFASNPPTTVSTNPSTPSPPPVQTOAANNST 120
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 119 SQEGLPAQ-APNTHDAAASLLSGRNVDTIINQLMEMGGGWDKQVORALRAAYNNPE 177
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 SSTDPTTVNSADTYGLAASNLVAGSNLEQTIQQIMDMGGGWDRTVSRALRAAYNNPE 180
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 178 RAVEVLYSGIPVTABIAVPVIGGGANTTDRAPTGBAGLSGIPNTAPLDLPFGQASNAGGG 237
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 RAIDLYSGIPAEAVAVPVPQTAGISSGAVPVG-----PNSPLANMFQ-ETXSCAG 232
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 238 AGGGPLDLRNNPQAVREVMVHTNPQILOPMLVLSKQNPQILRLIENHDEFQLLINE 297
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 233 AGLSGLDLRNNPQALRSMVQSNPQILQPVQLQELGKQNPSSLTLIQEHHAFLQLINE 292
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 298 PREGEGDFLQPEDEMPHAI SVTPQEATGRLESMPDRARVIEAPLACDRNEELAA 357
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 293 PVESGEGIDFQPEQD-MPHAINVTPAQEALGRLEAMGFORASVIEAPLACDHDEQLAA 351
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 358 NYLLEHAGE 366
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 352 NYLLENAGD 360
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 10

```
US-09-513-996A-18372
; Sequence 18372, Application US/09513996A
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: ENCODED THEREBY
; CURRENT APPLICATION NUMBER: US/09/513,996A
; CURRENT FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 81028
; SEQ ID NO 18372
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 1..371
; OTHER INFORMATION: any n or Xaa = unknown
; FEATURE:
; OTHER INFORMATION: Location 1..371 / Ceres Seq. ID 1396369
US-09-513-996A-18372
```

```
Query Match 61.2%; Score 1169.5; DB 19; Length 371;
Best Local Similarity 64.3%; Pred. No. 1.6e-87;
Matches 247; Conservative 43; Mismatches 65; Indels 29; Gaps 9;

Qy 1 MKLTVKTLKGTHTFEIRVQPNDTIMAVKKNIBEQKDSYPWGQQLLIPNGKVLKDESTLE 60
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MKLTVKTLKGSHTFEIRVLPSTDTIMAVKKNIBESQKDNYPGQQLLIHNGKVLKDESTLV 60
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 61 ENKVEDGLVWMLSGKTSSTGSSQHSN-----TPATRQAPLEAPQAPQPP 112
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 ENKVTIEGLVWMLSKSGSAGASVQTSVSVQPSVATSSTKPAAP--STTQSSPVP 118
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 113 VAPITTSQPEGLPAQAPNTHDAAASLLSGRNVDTIINQLMEMGGGWDKQVORALRAA 172
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 119 ASPIPAQEQ---PAAQTDYTGAASTLVSGSLEQMVQVQIMEMGGGWDKQVTRALRAA 175
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 173 YNNPERAVELYSGIPVTABIAVP-----IGGQANTTDRAPTGEAGLSGIPNTAPLDLP 227
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 176 YNNPERAVDLYSGIPQTAEVAVPPEAQIAGSGA-----APVAPA--SGGNSSPLDLF 228
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 228 PQGASNAGGAGGGLDLFLRNNPQFQAVREVMVHTNPQILOPMLVLSKQNPQILRLIEN 287
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 229 PQETVAAAGSGDLGLTLEFLRNNDFQQLRTVMVHSNPQILOPMLQELGKQNPQLRLIQEN 288
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 288 HDEFQLLINEPPEG--GEGDFLQPEDEMPHAI SVTPQEATGRLESMPDRARVIEA 345
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 289 QAEFLQLVNPEYEGSDGDMFQPEQ-EMPHAINVTPAQEALGRLEAMGFDRALVIEA 347
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 346 FLACDRNEELAAAYLLEHAGE-ED 368
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 348 FLACDRNEELAAAYLLENSGDFED 371
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 12

```
US-10-424-599-219691
; Sequence 219691, Application US/10424599
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
```

```
Db 289 QAEFLQLVNPEYEGSDGDMFQPEQ-EMPHAINVTPAQEALGRLEAMGFDRALVIEA 347
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 346 FLACDRNEELAAAYLLEHAGE-ED 368
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 348 FLACDRNEELAAAYLLENSGDFED 371
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 11
US-09-620-394B-6541
; Sequence 6541, Application US/09620394B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-1067P
; CURRENT APPLICATION NUMBER: US/09/620,394B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9131
; SEQ ID NO 6541
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..371
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: 1..371
; OTHER INFORMATION: Ceres Seq. ID 1396369
US-09-620-394B-6541
```

```
Query Match 61.2%; Score 1169.5; DB 20; Length 371;
Best Local Similarity 64.3%; Pred. No. 1.6e-87;
Matches 247; Conservative 43; Mismatches 65; Indels 29; Gaps 9;

Qy 1 MKLTVKTLKGTHTFEIRVQPNDTIMAVKKNIBEQKDSYPWGQQLLIPNGKVLKDESTLE 60
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MKLTVKTLKGSHTFEIRVLPSTDTIMAVKKNIBESQKDNYPGQQLLIHNGKVLKDESTLV 60
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 61 ENKVEDGLVWMLSGKTSSTGSSQHSN-----TPATRQAPLEAPQAPQPP 112
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 ENKVTIEGLVWMLSKSGSAGASVQTSVSVQPSVATSSTKPAAP--STTQSSPVP 118
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 113 VAPITTSQPEGLPAQAPNTHDAAASLLSGRNVDTIINQLMEMGGGWDKQVORALRAA 172
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 119 ASPIPAQEQ---PAAQTDYTGAASTLVSGSLEQMVQVQIMEMGGGWDKQVTRALRAA 175
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 173 YNNPERAVELYSGIPVTABIAVP-----IGGQANTTDRAPTGEAGLSGIPNTAPLDLP 227
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 176 YNNPERAVDLYSGIPQTAEVAVPPEAQIAGSGA-----APVAPA--SGGNSSPLDLF 228
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 228 PQGASNAGGAGGGLDLFLRNNPQFQAVREVMVHTNPQILOPMLVLSKQNPQILRLIEN 287
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 229 PQETVAAAGSGDLGLTLEFLRNNDFQQLRTVMVHSNPQILOPMLQELGKQNPQLRLIQEN 288
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 288 HDEFQLLINEPPEG--GEGDFLQPEDEMPHAI SVTPQEATGRLESMPDRARVIEA 345
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 289 QAEFLQLVNPEYEGSDGDMFQPEQ-EMPHAINVTPAQEALGRLEAMGFDRALVIEA 347
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 346 FLACDRNEELAAAYLLEHAGE-ED 368
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 348 FLACDRNEELAAAYLLENSGDFED 371
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 219691
LENGTH: 363
TYPE: PRT
ORGANISM: Glycine max
OTHER INFORMATION: Clone ID: PAT_MRT3847_40409C.1.pep
US-10-424-599-219691

Query Match 61.2%; Score 1168.5; DB 30; Length 363;
Best Local Similarity 63.1%; Pred. No. 1.9e-87;
Matches 236; Conservative 51; Mismatches 70; Indels 17; Gaps 6;

QY 1 MKLTVKTLKGSHFEIRVQPNQDITMAVKNIIEIQGKDSYPMQOQLLIHNGKVLKDESTLE 60
DB 1 MKLTVKTLKGSHFEIRVQPNQDITMAVKNIIEIQGKDSYPMQOQLLIHNGKVLKDESTLV 60

QY 61 ENKVNEDGFLVVLMSKGTSGTGTSSSOHSNTPAT-----RQAPLEAPQAPPPVAP 115
DB 61 ENKVNEDGFLVVLMSKGTSGTGTSSSOHSNTPAT-----RQAPLEAPQAPPPVAP 120

QY 116 ITTSOPEGLPAQPNTHDAAANLLSGRNVDITINQLMEMGGSGWDKQKQVQALRAAYNN 175
DB 121 SSTADP--TTNVSADTYGLAANLVAGSNLEQTIQIIMDMGGGNDRTVSRALRAAYNN 178

QY 176 PERAVEYLXGIPVTAEIAVPIG--GOGANTTDR--PTGEAGLSGIPNTAPDLDFPQASNAG 235
DB 179 PERAIDYLSGIPAEAAVVPVQTAGISSGAVPVG-----PNSPLMFPQ-ETISS 230

QY 236 GGAGGGLFLRNPFQAVREVMVHTNPQILQPMVLVELSKONPQILRLIEENHDFLOLL 295
DB 231 TGAGLSGLFLRNPFQALRSVQSNPQLQVQLQELGKQNPGLRLIQEHHGFLQI 290

QY 296 NEPFEGEGDFLDQPEDEMPHAI SVTPEQEAIGRLSMGDFDRARVIEAFACDRNEEL 355
DB 291 NEPFEGEGDFLDQPEDEMPHAI SVTPEQEAIGRLSMGDFDRARVIEAFACDRNEEL 349

QY 356 AANYLLEHAGE-ED 368
DB 350 AANYLLENAGDFED 363

RESULT 13
US-09-791-537-55742
Sequence 55742, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
TITLE OF INVENTION: Danzer, Joseph
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 55742
LENGTH: 379
TYPE: PRT
ORGANISM: Daucus carota
US-09-791-537-55742

Query Match 60.4%; Score 1153.5; DB 22; Length 379;
Best Local Similarity 63.0%; Pred. No. 3.5e-86;
Matches 240; Conservative 45; Mismatches 81; Indels 15; Gaps 9;

QY 1 MKLTVKTLKGTHFEIRVQPNQDITMAVKNIIEIQGKDSYPMQOQLLIHNGKVLKDESTLE 60

DB 1 MKLTVKTLKGSHFEIRVQPNQDITMAVKNIIEIQGKDSYPMQOQLLIHNGKVLKDESTLA 60
QY 61 ENKVNEDGFLVVLMSKGTSGTGTSSSOHSNTPATQAP---PLEAPQAPQPPVAPIT 117
DB 61 ESKISEDGFLVVLMSKGTSGTGTSSSOHSNTPATQAP---PLEAPQAPQPPVAPIT 120

QY 118 TSQPEG--LPAQAP-NTHDAAANLLSGRNVDITINQLMEMGGSGWDKQKQVQALRAAYNN 174
DB 121 T-VPEAPLSFAFAPSDTYGEAASNVAGSNLEQTIQIIMDMGGGNDRTVSRALRAAYNN 179

QY 175 NPERAVEYLXGIPVTAEIAVPIG--GOGANTTDR--PTGEAGLSGIPNTAPDLDFP 228
DB 180 NPERAVEYLXGIPVTAEIAVPIG--GOGANTTDR--PTGEAGLSGIPNTAPDLDFP 239

QY 229 QGASNAGGAGGGLFLRNPFQAVREVMVHTNPQILQPMVLVELSKONPQILRLIEENH 288
DB 240 QETLSGVTGAGLSGLFLRNPFQAVREVMVHTNPQILQPMVLVELSKONPQILRLIEENH 299

QY 289 DEFLQLLNEPPEGEGDFLDQPEDEMPHAI SVTPEQEAIGRLSMGDFDRARVIEAFLA 348
DB 300 EEFQLQINEPVEASEGDMFQPEQD-VPEQETVTAADQEAERLEAMGDFDRGLVIEAFLA 358

QY 349 CDRNEELAANYLLEHAGE-ED 368
DB 359 CDRNEELAVNYLLENAGDFED 379

RESULT 14
US-09-513-996A-55531
Sequence 55531, Application US/09513996A
GENERAL INFORMATION:
APPLICANT: N. ALEXANDROV et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDE
FILE REFERENCE: 2750-709P
CURRENT APPLICATION NUMBER: US/09/513,996A
CURRENT FILING DATE: 2000-02-25
NUMBER OF SEQ ID NOS: 81028
SEQ ID NO 55531
LENGTH: 368
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: UNSURE
LOCATION: 1..368
OTHER INFORMATION: any n or Xaa = unknown
FEATURE:
OTHER INFORMATION: Location 1..368 / Ceres Seq. ID 2113785
US-09-513-996A-55531

Query Match 58.4%; Score 1115; DB 19; Length 368;
Best Local Similarity 62.8%; Pred. No. 5.1e-83;
Matches 235; Conservative 45; Mismatches 82; Indels 12; Gaps 8;

QY 1 MKLTVKTLKGTHFEIRVQPNQDITMAVKNIIEIQGKDSYPMQOQLLIHNGKVLKDESTLE 60
DB 1 MKLTVKTLKGSHFEIRVLPDITMAVKNIIEIQGKDSYPMQOQLLIHNGKVLKDESTLV 60

QY 61 ENKVNEDGFLVVLMSKGTSGTGTSSSO-HSNTPATQAPLEAPQAPQPPVAPITTS 119
DB 61 ENKVNEDGFLVVLMSKGTSGTGTSSSO-HSNTPATQAPLEAPQAPQPPVAPITTS 119

QY 120 QP-EGLEPAQPNTHDAAANLLSGRNVDITINQLMEMGGSGWDKQKQVQALRAAYNN 178
DB 120 TPVQEQPTAQSDTYGQAASLVSGSSIEQMVQOIMEMGGSGWDKETVTRALRAAYNN 179

QY 179 AVEYLXGIPVTAEI-AVPIGOGANTTDRPTGEAGLSGIPNTAPDLDFPQASNAGG 237
DB 180 AVDYLSGIPETVTPATNLVSGSGRELTAPPP-----SGGNSSPLDLFPQASVDAAG 235

QY 238 AGGGLDFLRNPFQAVREVMVHTNPQILQPMVLVELSKONPQILRLIEENHDFLOLLNE 297

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 06:16:40 ; Search time 13.3299 Seconds
(without alignments)
1168.082 Million cell updates/sec

Title: US-09-805-550-4

Perfect score: 1910

Sequence: 1 MKLTVKTLKGTHTPEIRVQPN.....CDRNEELAANYLLEHAGEED 368

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1910	100.0	368	3	US-09-413-574-4
2	900.5	47.1	405	3	US-09-413-574-2
3	182	9.5	79	3	US-09-100-802-4
4	173.5	9.1	126	4	US-08-747-562-26
5	164	8.6	81	3	US-09-100-802-5
6	154	8.1	915	2	US-08-480-917-2
7	154	8.1	915	3	US-09-138-736-2
8	154	8.1	915	4	US-08-988-242-2
9	133.5	7.0	1274	4	US-09-095-443-2
10	126.5	6.6	156	2	US-09-070-060-7
11	126.5	6.6	156	3	US-09-051-969A-3
12	126.5	6.6	156	3	US-09-051-969A-4
13	126.5	6.6	156	3	US-09-357-746-7
14	126.5	6.6	160	4	US-09-370-838-205
15	126.5	6.6	229	2	US-08-726-306A-23
16	126.5	6.6	229	3	US-08-840-146-20
17	126.5	6.6	229	3	US-09-360-220-20
18	126	6.6	2616	6	5206163-3
19	125.5	6.6	198	4	US-08-988-242-19
20	125.5	6.6	352	3	US-08-854-764-2
21	125.5	6.6	352	5	PCT-US95-09377-2
22	123.5	6.5	76	3	US-08-817-787-8
23	123.5	6.4	76	2	US-09-070-060-8
24	122.5	6.4	76	3	US-09-357-746-8
25	122.5	6.4	76	3	US-08-817-787-1
26	122.5	6.4	76	3	US-09-100-802-1
27	122.5	6.4	76	4	US-09-331-930A-26

28 122.5 6.4 103 2 US-08-771-201-9 Sequence 9, Appli
29 122.5 6.4 114 2 US-08-771-201-10 Sequence 10, Appl
30 122.5 6.4 147 2 US-08-771-201-11 Sequence 11, Appl
31 121 6.3 1121 1 US-07-789-915A-2 Sequence 2, Appli
32 121 6.3 1121 1 US-08-005-002C-2 Sequence 2, Appli
33 121 6.3 1121 1 US-08-487-203A-2 Sequence 2, Appli
34 119.5 6.3 323 2 US-08-747-788-2 Sequence 2, Appli
35 119.5 6.3 323 4 US-09-300-681B-2 Sequence 2, Appli
36 117.5 6.2 457 4 US-09-516-277-2 Sequence 2, Appli
37 117.5 6.2 457 4 US-09-516-277-4 Sequence 2, Appli
38 117.5 6.2 533 1 US-08-462-092-2 Sequence 4, Appli
39 117.5 6.2 533 3 US-08-746-822-2 Sequence 2, Appli
40 117.5 6.2 533 3 US-09-094-350-2 Sequence 2, Appli
41 117.5 6.2 533 6 5510474-2 Patent No. 5510474
42 116 6.1 447 4 US-09-480-921B-6 Sequence 6, Appli
43 115 6.0 158 4 US-09-091-725-25 Sequence 25, Appl
44 114.5 6.0 76 1 US-08-232-815-2 Sequence 2, Appli
45 114.5 6.0 76 1 US-08-350-906-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-413-574-4
; Sequence 4, Application US/09413574
; Patent No. 6235972
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; APPLICANT: Tegliani, Laura
; TITLE OF INVENTION: Maize Rad23 Genes and Uses Thereof
; FILE REFERENCE: 0964
; CURRENT APPLICATION NUMBER: US/09/413,574
; CURRENT FILING DATE: 1999-10-06
; EARLIER APPLICATION NUMBER: 60/109,728
; EARLIER FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Zea mays
US-09-413-574-4

Query Match 100.0%; Score 1910; DB 3; Length 368;
Best Local Similarity 100.0%; Pred. No. 8.3e-164;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MKLTVKTLKGTHTPEIRVQPNNTIMAVKKNIEEIQGKDSYPMGQOLLIENGKVLKDESTLE 60
Db 1 MKLTVKTLKGTHTPEIRVQPNNTIMAVKKNIEEIQGKDSYPMGQOLLIENGKVLKDESTLE 60
Qy 61 ENKVEDGFLVWLSKGTSGTGTSSQHSNTPATROAPPLPAPPOAPPPVAPITTSQ 120
Db 61 ENKVEDGFLVWLSKGTSGTGTSSQHSNTPATROAPPLPAPPOAPPPVAPITTSQ 120
Qy 121 PEGLPQAQPNTHNAASNLSSGRNVTIINQLMEMGGGSKDKVORALRAAYNNPERAV 180
Db 121 PEGLPQAQPNTHNAASNLSSGRNVTIINQLMEMGGGSKDKVORALRAAYNNPERAV 180
Qy 181 EYLISGIPVTAETAVPIGGQANTTDRAPTGEAGLSIPNTAPLDLFPQASNAGGAGG 240
Db 181 EYLISGIPVTAETAVPIGGQANTTDRAPTGEAGLSIPNTAPLDLFPQASNAGGAGG 240
Qy 241 GPLDFLRNNPQOAVREVMVHTNPQILQPMVLVELSKQNPQILRIENHDEFLQLNEPPE 300
Db 241 GPLDFLRNNPQOAVREVMVHTNPQILQPMVLVELSKQNPQILRIENHDEFLQLNEPPE 300
Qy 301 GEGDGLDQPEEDEMHPHAI SVTPEEQEATGLESMGFDRARVIEAFLACDRNEELANYL 360
Db 301 GEGDGLDQPEEDEMHPHAI SVTPEEQEATGLESMGFDRARVIEAFLACDRNEELANYL 360
Qy 361 LEHAGED 368

Wed Dec 17 06:49:42 2003

```
Db 361 LEHAGEED 368
|||||
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
US-09-100-802-4

Query Match 9.5%; Score 182; DB 3; Length 79;
Best Local Similarity 43.6%; Pred. No. 2.9e-09;
Matches 34; Conservative 23; Mismatches 21; Indels 0; Gaps 0;

Qy 1 MKLTVKTLKGTHTFEIRVQPNDDTMAVKNIIEIQKDSYPMGQQLLIIFNGKVLKDESTLE 60
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1 MQVTLKTLQQTQFKIDIDPEETVKALKERIESEKGDAPFVAGQKLIYAGKILMDTALK 60
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 61 ENKYNEDGFLVWMLSGK 78
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 61 EYKIDENFVVVMVTKPK 78
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 4
US-08-747-562-26
; Sequence 26 Application US/08747562
; Patent No. 6579897
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: BOLDIN, Mark
; APPLICANT: METT, Igor
; APPLICANT: VARFOLOMEV, Eugene
; TITLE OF INVENTION: MODULATOR OF TNF/NGF SUPERFAMILY RECEPTORS
; TITLE OF INVENTION: AND SOLUBLE OLIGOMERIC TNF/NGF SUPERFAMILY RECEPTORS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/747,562
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05854
; FILING DATE: 11-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 109,632
; FILING DATE: 11-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 111,125
; FILING DATE: 02-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: WALLACH=15A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-747-562-26

Db 361 LEHAGEED 368
|||||
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Zea mays
US-09-413-574-2

Query Match 47.1%; Score 900.5; DB 3; Length 405;
Best Local Similarity 49.0%; Pred. No. 7e-73;
Matches 197; Conservative 53; Mismatches 117; Indels 35; Gaps 7;

Qy 1 MKLTVKTLKGTHTFEIRVQPNDDTMAVKNIIEIQKDSYPMGQQLLIIFNGKVLKDESTLE 60
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1 MKLNVKTLKGTNFEIEASPDASVADVKRIETTTQGSTYRADOQMLLYQGKILKDETTLE 60
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 61 ENKYNEDGFLVWMLSGKTSSTSSOHSNTPTATQAP-----PLEAPQAP-QPPVA 114
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 61 SNGVAENSLVLMLSKAKAS-SGGASTATTAKAPATLAQPAAPVAPAAASVARTPTQAPVA 119
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 115 PITTSOPEGLPAQAP-----NTHDNAASNLISGRNVDTIINQLMEMGGSGWDKDK 164
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 120 TAEATPSPVQPAAPAAATVAATDDADVYQASNLVFGNLSQTLQIILDMGGTWERDT 179
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 165 VQALRAAYNNPERAVEYIIGIPVTAE-----IAPVIGGGANTTDRAPTGEA----- 213
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 180 VVRLAAYNNPERAIDYLSGIPENVEAPQVARAPAAQQTNOQAASPAQPAVALPVQP 239
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 214 -GLSGIPNTAPLDFPQASNAGG-----GAGGGLDFLRNPNPOQAVREMYHTNPQIL 266
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 240 SPASAGFNANPLNLPFGVPSGSGNPGVPGAGSGALDALQLPQFQALLQLVQANPQIL 299
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 267 QPMLVELSKONPOILRLIENHDEFLQLNEPPEGGEGDFLDQPEDEMPHAIISVTTPERQ 326
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 300 QPMLQELGKONPOILRLIQENQAEFLRLVNESPEGGPGGNILQOLAAAVPQTLVTTPER 359
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 327 EAIGRLSMGFDRARVIEAFACDRNEELAAVYLLHEAGEED 368
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 360 EAIGRLSMGFDRARVIEAFACDRNEELAAVYLLHEAGEED 401
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 3
US-09-100-802-4
; Sequence 4, Application US/09100802A
; Patent No. 6294363
; GENERAL INFORMATION:
; APPLICANT: Madura, Kiran
; TITLE OF INVENTION: Methods and Compositions for the Rapid
; TITLE OF INVENTION: Purification of Proteasomes and Methods of Use of Components
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: UMDNJ97-11
; CURRENT APPLICATION NUMBER: US/09/100,802A
; CURRENT FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/050,171
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 17
```


; APPLICATION NUMBER: US/08/988,242
; FILING DATE: 10-DEC-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 36400A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 915 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: /desc = "protein"
; US-08-988-242-2

Query Match 8.1%; Score 154; DB 4; Length 915;
Best Local Similarity 22.9%; Pred. No. 3.6e-05;
Matches 56; Conservative 43; Mismatches 126; Indels 20; Gaps 6;

QY 57 STLEENKYNEDGLVLMVLSKGTSGTSSQHSNTPATROAPLEAQQAPQPPVAPI 116
Db 356 SRREEQPGQKTSVVATAKPGCVSSGTDASSSHNTTSAASASPSPVSAKAAAP- 414
QY 117 TTSQEGPLPAQANTHDNAASNLISGRNVDTIINQLEMGGGSDKDKVQALRAAYNPP 176
Db 415 ----PAAARSAPHVGSKIIANLVNQLGINTVQSVSTGAPATTR---STAVTSTTTAP 467
QY 177 ERAVELYSGIPVTAETIAVPIGGGANTTDRA---PTGEAGLSGIPNTAPLDLFPQASN 233
Db 468 QRTSPYGHNGRPVTAGIVAANSASAASPTAAAKPTGEKASACETSSVAINATRPAL 527
QY 234 AGGAGGGPLDPLRNPPQFQAVREMVHTNPQIQLPMLVELSKNQPIRL-----IEENHD 289
Db 528 HNASLPQAPTDGVLAAAVYQSEGE-VHQSLESLVESVITNTSR---VLKLLPDTIRRDHE 582
QY 290 EFLQL 294
Db 583 QLLNL 587

RESULT 9
US-09-095-443-2
; Sequence 2, Application US/09095443
; Patent No. 6342593
; GENERAL INFORMATION:
; APPLICANT: Plowman, Gregory
; APPLICANT: Peles, Eior
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF ALP RELATED DISORDERS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,443
; FILING DATE: Herewith
; CLASSIFICATION:

; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/480,917
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 36400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 915 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "protein"
; US-09-138-736-2

Query Match 8.1%; Score 154; DB 3; Length 915;
Best Local Similarity 22.9%; Pred. No. 3.6e-05;
Matches 56; Conservative 43; Mismatches 126; Indels 20; Gaps 6;

QY 57 STLEENKYNEDGLVLMVLSKGTSGTSSQHSNTPATROAPLEAQQAPQPPVAPI 116
Db 356 SRREEQPGQKTSVVATAKPGCVSSGTDASSSHNTTSAASASPSPVSAKAAAP- 414
QY 117 TTSQEGPLPAQANTHDNAASNLISGRNVDTIINQLEMGGGSDKDKVQALRAAYNPP 176
Db 415 ----PAAARSAPHVGSKIIANLVNQLGINTVQSVSTGAPATTR---STAVTSTTTAP 467
QY 177 ERAVELYSGIPVTAETIAVPIGGGANTTDRA---PTGEAGLSGIPNTAPLDLFPQASN 233
Db 468 QRTSPYGHNGRPVTAGIVAANSASAASPTAAAKPTGEKASACETSSVAINATRPAL 527
QY 234 AGGAGGGPLDPLRNPPQFQAVREMVHTNPQIQLPMLVELSKNQPIRL-----IEENHD 289
Db 528 HNASLPQAPTDGVLAAAVYQSEGE-VHQSLESLVESVITNTSR---VLKLLPDTIRRDHE 582
QY 290 EFLQL 294
Db 583 QLLNL 587

RESULT 8
US-08-988-242-2
; Sequence 2, Application US/08988242
; Patent No. 6403103
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: LESENECHAL, MYLENE
; APPLICANT: JOLIVET, MICHEL
; APPLICANT: MANBRAND, BERNARD
; TITLE OF INVENTION: NEW TRYPAANOSOMA CRUZI ANTIGEN, GENE
; TITLE OF INVENTION: ENCODING THEREFOR, AND METHODS OF DETECTING AND TREATING
; TITLE OF INVENTION: CHAGAS DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

1 FILING DATE: 30-APR-1998
2 CLASSIFICATION:
3 PRIOR APPLICATION DATA:
4 APPLICATION NUMBER: 60/073,839
5 FILING DATE: 05-FEB-1998
6 ATTORNEY/AGENT INFORMATION:
7 NAME: Higgins, Patrick H
8 REGISTRATION NUMBER: 39,709
9 REFERENCE/DOCKET NUMBER: PHM.70312
10 TELECOMMUNICATION INFORMATION:
11 TELEPHONE: 302.886.4889
12 TELEFAX: 302.886.8221
13 TELEX:
14 INFORMATION FOR SEQ ID NO: 7:
15 SEQUENCE CHARACTERISTICS:
16 LENGTH: 156 amino acids
17 TYPE: amino acid
18 STRANDEDNESS: unknown
19 TOPOLOGY: unknown
20 MOLECULE TYPE: peptide
21 US-09-070-060-7
22
23 Query Match 6.6%; Score 126.5; DB 2; Length 156;
24 Best Local Similarity 35.4%; Pred. No. 0.00078;
25 Matches 34; Conservative 16; Mismatches 37; Indels 9; Gaps 2;
26
27 QY 1 MKLVTKTKGTFEIRVOPNDTMAVKNTIEEIQKDSYPWGQOLLIFNGKVLKDESTLE 60
28 | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
29 Db 1 MQIFVKLTGKTTITLEVPSDTTENVK---AKIQKEGIPDQORLIFAGKQLEDGRTL 57
30 | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
31 QY 61 ENKVNEDGFLVVMVLM-----SKGTSGSTGTSSSQH 90
32 | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
33 Db 58 DYNIQKESTLHLVLRGGAKKKKSYTTPKKNKH 93
34 | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
35
36 RESULT 11
37 US-09-051-969A-3
38 ; Sequence 3, Application US/09051969A
39 ; Patent No. 6063598
40 ; GENERAL INFORMATION:
41 ; APPLICANT: ENENKEL, BARBARA
42 ; APPLICANT: GANNON, FRANK
43 ; APPLICANT: BERGEMANN, KLAUS
44 ; APPLICANT: NOE, WOLFGANG
45 ; TITLE OF INVENTION: INTENSIVE HOMOLOGOUS PROMOTER OBTAINED
46 ; TITLE OF INVENTION: FROM HAMSTERS
47 ; NUMBER OF SEQUENCES: 8
48 ; CORRESPONDENCE ADDRESS:
49 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
50 ; STREET: 1100 NEW YORK AVENUE, SUITE 600
51 ; CITY: WASHINGTON
52 ; STATE: DC
53 ; COUNTRY: USA
54 ; ZIP: 20005
55 ; COMPUTER READABLE FORM:
56 ; MEDIUM TYPE: Floppy disk
57 ; COMPUTER: IBM PC compatible
58 ; OPERATING SYSTEM: PC-DOS/MS-DOS
59 ; SOFTWARE: PatencIn Release #1.0, Version #1.30
60 ; CURRENT APPLICATION DATA:
61 ; APPLICATION NUMBER: US/09/051,969A
62 ; FILING DATE: 1998-09-30
63 ; CLASSIFICATION: 435
64 ; ATTORNEY/AGENT INFORMATION:
65 ; NAME: FLESHNER, RAZ E.
66 ; REGISTRATION NUMBER: 34,331
67 ; REFERENCE/DOCKET NUMBER: 0652.1690000
68 ; TELECOMMUNICATION INFORMATION:
69 ; TELEPHONE: 202-371-2600
70 ; TELEFAX: 202-371-2540
71 ; INFORMATION FOR SEQ ID NO: 3:
72 ; SEQUENCE CHARACTERISTICS:
73 ; LENGTH: 156 amino acids

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 06:27:31 ; Search time 18.3376 Seconds
(without alignments)
2123.954 Million cell updates/sec

Title: US-09-805-550-2

Perfect score: 405

Sequence: 1 MKNLVKTKGNFIEASPD.....BELTANYLLDHGHEFDQQQ 405

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168882 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

PIR_76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	4.4	113	2 H86296	F309.1 protein - A
2	17	4.2	382	2 T14336	RAD23 protein, iso
3	14	3.5	392	2 T04150	RAD23 protein homo
4	13	3.2	246	2 G86296	T24D18.27 protein
5	13	3.2	367	2 F96827	protein F20B17.8 [
6	13	3.2	379	2 T14337	RAD23 protein, iso
7	9	2.2	3069	2 H70656	fatty-acid synthas
8	8	2.0	61	2 B61384	Cowper's gland muc
9	8	2.0	254	1 BVECA	pyruvate dehydroge
10	8	2.0	254	2 E85494	pyruvate dehydroge
11	8	2.0	254	2 E90643	pyruvate dehydroge
12	8	2.0	254	2 AF0415	pyruvate dehydroge
13	8	2.0	254	2 AF0521	pyruvate dehydroge
14	8	2.0	256	2 E82079	pyruvate dehydroge
15	8	2.0	266	2 B33034	streptomycin resis
16	8	2.0	273	2 T03254	probable carbonate
17	8	2.0	289	2 A82006	2'-hydroxyisoflavo
18	8	2.0	300	2 T08106	hypothetical prote
19	8	2.0	339	2 T34406	cellulase egli - s
20	8	2.0	393	2 S59499	photolase/cryptoc
21	8	2.0	462	2 A84327	high affinity sulf
22	8	2.0	566	2 D74633	probable membran
23	8	2.0	643	2 D70697	transposase BME112
24	8	2.0	653	2 A13404	general secretion
25	8	2.0	687	2 D87270	DNA-directed DNA p
26	8	2.0	1235	1 DUBEV1	DNA-directed DNA p
27	8	2.0	1235	1 DUBEH7	FmtB protein [impo
28	8	2.0	2481	2 D90011	hypothetical prote
29	7	1.7	58	2 C84025	

30	7	1.7	87	2	AG1930	hypothetical prote
31	7	1.7	100	2	B89979	conserved hypothet
32	7	1.7	108	2	AG0699	probable pathogeni
33	7	1.7	128	2	G91250	hypothetical prote
34	7	1.7	129	2	T36732	hypothetical prote
35	7	1.7	139	2	B86810	transcription regu
36	7	1.7	144	2	S53430	adenyde dehydroge
37	7	1.7	144	2	AB2443	hypothetical prote
38	7	1.7	154	2	D86749	riboflavin synthas
39	7	1.7	163	2	C90281	hypothetical prote
40	7	1.7	176	2	F75363	hypothetical prote
41	7	1.7	176	2	E95274	hypothetical prote
42	7	1.7	183	2	F71304	probable ribosome
43	7	1.7	188	2	I51347	MHC class I - Atla
44	7	1.7	218	2	H84364	membrane protein [
45	7	1.7	226	2	T29210	hypothetical prote
46	7	1.7	227	2	T45388	hypothetical prote
47	7	1.7	228	2	S46965	microfilarial shea
48	7	1.7	231	2	E86338	protein F5M15.16 [
49	7	1.7	231	2	A13283	transcription regu
50	7	1.7	232	2	AF3406	hypothetical expor
51	7	1.7	238	2	H83559	hypothetical prote
52	7	1.7	242	2	E70822	hypothetical prote
53	7	1.7	245	2	AD0329	conserved hypothet
54	7	1.7	245	2	AC0830	conserved hypothet
55	7	1.7	247	2	T50874	hypothetical cytoc
56	7	1.7	252	2	JC5237	osmotin-like prote
57	7	1.7	252	2	T29527	hypothetical prote
58	7	1.7	254	2	C82868	plasmid replicatio
59	7	1.7	254	2	T25073	hypothetical prote
60	7	1.7	256	2	S76756	1-(5-phosphoribosy
61	7	1.7	256	2	B32393	T-cell antigen 4-1
62	7	1.7	257	2	T46548	probable transcrip
63	7	1.7	258	2	A75618	probable chromosom
64	7	1.7	264	2	G83245	probable ATP-bindi
65	7	1.7	266	2	T10644	hypothetical prote
66	7	1.7	280	2	S72844	beta-aspartate met
67	7	1.7	280	2	A70513	hypothetical prote
68	7	1.7	280	2	A61047	ectodermal (ect) -
69	7	1.7	281	2	C87073	conserved hypothet
70	7	1.7	281	2	A43749	lignin beta-ether
71	7	1.7	290	2	AG3297	biopolymer transpo
72	7	1.7	292	2	T18584	F5C11.2 - Caenorh
73	7	1.7	301	2	C96961	cell division prot
74	7	1.7	303	2	F81318	probable galactosy
75	7	1.7	303	2	F87629	hypothetical prote
76	7	1.7	307	2	C75584	hypothetical prote
77	7	1.7	315	2	T03733	hypothetical prote
78	7	1.7	322	2	F88029	protein F46F5.5 [i
79	7	1.7	324	2	T04478	probable carbonate
80	7	1.7	330	2	G84127	phosphotransacetyl
81	7	1.7	335	2	H84170	glyceraldehyde 3-p
82	7	1.7	338	2	H87510	polyprenyl synthet
83	7	1.7	338	2	S42594	hypothetical prote
84	7	1.7	340	2	A82218	transcription regu
85	7	1.7	343	2	I51348	MHC class I - Atla
86	7	1.7	347	2	T31755	hypothetical prote
87	7	1.7	351	2	T33976	hypothetical prote
88	7	1.7	358	2	S44233	strp protein - Str
89	7	1.7	364	1	BVEECIC	incC protein - pla
90	7	1.7	365	2	I84432	MHC class I protei
91	7	1.7	366	2	JU0469	type II site-speci
92	7	1.7	378	2	H82640	conserved hypothet
93	7	1.7	385	2	G71467	probable succinyl-
94	7	1.7	387	2	T30241	hypothetical prote
95	7	1.7	393	2	E82283	conserved hypothet
96	7	1.7	396	2	A82532	translation elonga
97	7	1.7	400	1	QREBVT	glycine betaine/pr
98	7	1.7	400	1	BVECPV	hypothetical prote
99	7	1.7	400	2	D91071	hypothetical prote
100	7	1.7	400	2	G85915	

ALIGNMENTS

```

RESULT 1
H86296
C:Species: Arabidopsis thaliana
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
C:Accession: H86296
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: H86296
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-113 <STO>
A:Cross-references: GB:AE005172; NID:g4966345; PIDN:AD34676.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 4.4%; Score 18; DB 2; Length 113;
Best Local Similarity 100.0%; Pred. No. 8.4e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 295 NPQILQMLQELGKQNPQ 312
Db 5 NPQILQMLQELGKQNPQ 22

RESULT 2
T14336
RAD23 protein, isoform I - carrot
C:Species: Daucus carota (carrot)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T14336
R:Sturm, A.; Leinhardt, S.
Plant J. 13, 815-821, 1998
A:Title: Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in yeast.
A:Reference number: Z17989; MUID:98345997; PMID:9681019
A:Accession: T14336
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-382 <STO>
A:Cross-references: EMBL:Y12013; NID:g1914682; PIDN:CAA72741.1; PID:g1914683
A:Experimental source: subspecies Queen Anne's Lace, isolate W001C
C:Genetics:
A:Gene: RAD23-1
C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

Query Match 4.2%; Score 17; DB 2; Length 382;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 295 NPQILQMLQELGKQNPQ 311
Db 282 NPQILQMLQELGKQNPQ 298

RESULT 3
T04150
RAD23 protein homolog - rice
C:Species: Oryza sativa (rice)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: T04150
R:Schultz, T.F.; Quatrano, R.S.

```

```

Plant Mol. Biol. 34, 557-562, 1997
A:Title: Characterization and expression of a rice RAD23 gene.
A:Reference number: Z08695; MUID:97369378; PMID:9225866
A:Accession: T04150
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-392 <SCH>
A:Cross-references: EMBL:U63530; NID:g1488296; PIDN:AAB65841.1; PID:g1488297
A:Experimental source: cv. Nipponbare
C:Genetics:
A:Gene: RAD23
C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

Query Match 3.5%; Score 14; DB 2; Length 392;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 VADVKRIIETTQGG 36
Db 23 VADVKRIIETTQGG 36

RESULT 4
G86296
T24D18.27 protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
C:Accession: G86296
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallor
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: G86296
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-246 <STO>
A:Cross-references: GB:AE005172; NID:g6587822; PIDN:AAF18513.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 3.2%; Score 13; DB 2; Length 246;
Best Local Similarity 100.0%; Pred. No. 0.00014;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 182 RALRAAYNNPERA 194
Db 155 RALRAAYNNPERA 167

RESULT 5
F96827
protein F20B17.8 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002
C:Accession: F96827
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallor
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712

```

A:Accession: F96827
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-367 <STO>
A:Cross-references: GB:AE005173; NID:g7715605; PIDN:AAF68123.1; GSPDB:GN00141
C:Genetics:
A:Gene: F20B17.8
A:Map position: 1
C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

Query Match 3.2%; Score 13; DB 2; Length 367;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 182 RALRAAYNNPERA 194
|||||
DB 168 RALRAAYNNPERA 180
|||||

RESULT 6
T14337
RAD23 protein, isoform II - carrot
C:Species: Daucus carota (carrot)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T14337
R:Sturm, A.; Leinhardt, S.
Plant J. 13, 815-821, 1998
A:Title: Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in yeast.
A:Reference number: Z17989; MUID:98345997; PMID:9681019
A:Accession: T14337
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-379 <STU>
A:Cross-references: EMBL:Y12014; NID:g1914684; PIDN:CAAT72742.1; PID:g1914685
A:Experimental source: subspecies Queen Anne's Lace, isolate W001C
C:Genetics:
A:Gene: RAD23-2
C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

Query Match 3.2%; Score 13; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 182 RALRAAYNNPERA 194
|||||
DB 172 RALRAAYNNPERA 184
|||||

RESULT 7
H70656
fatty-acid synthase (EC 2.3.1.85) - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: H70656
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: H70656
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-3069 <COL>
A:Cross-references: GB:Z83863; GB:AL123456; NID:g3261685; PIDN:CAB06201.1; PID:g3261686
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: fas
C:Superfamily: Brevibacterium ammoniagenes fatty-acid synthase
C:Keywords: acyltransferase; coenzyme A

Query Match 2.2%; Score 9; DB 2; Length 3069;

Best Local Similarity 100.0%; Pred. No. 12;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 AAPVAPAAAS 108
|||||
DB 1754 AAPVAPAAAS 1762
|||||

RESULT 8
B61384
Cowper's gland mucin glycoprotein - pig (fragments)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 07-Feb-1997
C:Accession: B61384
R:Sangadala, S.; Kim, D.; Brewer, J.M.; Mendicino, J.
Mol. Cell. Biochem. 102, 71-93, 1991
A:Title: Subunit structure of deglycosylated human and swine trachea and Cowper's gland
A:Reference number: A61384; MUID:91270244; PMID:2052001
A:Accession: B61384
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-61 <SAN>
C:Keywords: glycoprotein

Query Match 2.0%; Score 8; DB 2; Length 61;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 239 PSPASAGP 246
|||||
DB 12 PSPASAGP 19
|||||

RESULT 9
BVECA
pyruvate dehydrogenase complex repressor - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 31-Mar-1989 #sequence_revision 05-Dec-1997 #text_change 01-Mar-2002
C:Accession: A64734; A30263; S45192; I53482
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; .A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: A64734
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-254 <BLAT>
A:Cross-references: GB:AE000120; GB:U00096; NID:g1786298; PIDN:AAC73224.1; PID:g178630
A:Experimental source: strain K-12, substrain MG1655
R:Stephens, P.E.; Darlison, M.G.; Lewis, H.M.; Guest, J.R.
Eur. J. Biochem. 133, 153-162, 1983
A:Title: The pyruvate dehydrogenase complex of Escherichia coli K12. Nucleotide sequen
A:Reference number: A91130; MUID:83209630; PMID:6343085
A:Accession: A30263
A:Molecule type: DNA
A:Residues: 1-229, 'DFARQK', 236 <STE>
A:Cross-references: EMBL:V01498
A:Note: the gene for this protein, called gene A by the authors, is situated upstream
R:Fujita, N.
submitted to the EMBL Data Library, January 1994
A:Reference number: S45181
A:Accession: S45192
A:Molecule type: DNA
A:Residues: 1-229, 'DFARQK', 236 <FU>
A:Cross-references: EMBL:D26562; NID:g473770; PIDN:BAA05571.1; PID:g473782
A:Experimental source: strain K-12 substrain W3110
R:Haydon, D.J.; Quail, M.A.; Guest, J.R.
FEBS Lett. 336, 43-47, 1993
A:Title: A mutation causing constitutive synthesis of the pyruvate dehydrogenase compl.
A:Reference number: I53482; MUID:94085588; PMID:8262214
A:Accession: I53482
A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: DNA
 A:Residues: 1-117, 'C', 119-254 <RES>
 A:Cross-references: GB:S67363; NID:9455921; PIDN:AB29356.1; PID:9455922
 C:Genetics:
 A:Gene: pdhR
 A:Map position: 3 min
 C:Superfamily: pyruvate dehydrogenase complex repressor
 C:Keywords: DNA binding, transcription regulation

Query Match 2.0%; Score 8; DB 1; Length 254;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 REAIQRL 366
 |||||
 DB 52 REAIQRL 59

RESULT 10

E85494
 Pyruvate dehydrogenase complex repressor - Escherichia coli (strain O157:H7, substrain E85494)
 C:Species: Escherichia coli
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 19-Jul-2002
 C:Accession: E85494
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; NID:21074935; PMID:11206551
 A:Accession: E85494
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-254 <STO>
 A:Cross-references: GB:AE005174; NID:912512822; PIDN:AAG54417.1; GSPDB:GN00145; UMGF:Z01
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: pdhR
 C:Superfamily: pyruvate dehydrogenase complex repressor

Query Match 2.0%; Score 8; DB 2; Length 254;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 REAIQRL 366
 |||||
 DB 52 REAIQRL 59

RESULT 11

E90643
 Pyruvate dehydrogenase complex repressor - Escherichia coli (strain O157:H7, substrain E90643)
 C:Species: Escherichia coli
 C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 19-Jul-2002
 C:Accession: E90643
 R:Havashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
 A:Reference number: A99629; NID:21156231; PMID:11258796
 A:Accession: E90643
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-254 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA833540.1; PID:gl3359573; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RMD 0509952
 C:Genetics:
 A:Gene: ECs0117
 C:Superfamily: pyruvate dehydrogenase complex repressor

Query Match 2.0%; Score 8; DB 2; Length 254;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 REAIQRL 366
 |||||
 DB 52 REAIQRL 59

RESULT 12

AF0415
 Pyruvate dehydrogenase complex repressor [imported] - Yersinia pestis (strain CO92)
 C:Species: Yersinia pestis
 C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 19-Jul-2002
 C:Accession: AF0415
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.
 il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell
 Nature 413, 523-527, 2001
 A>Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AF0415
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-254 <KUR>
 A:Cross-references: GB:AL590842; PIDN:CAC92650.1; PID:gl5981347; GSPDB:GN00175
 C:Genetics:
 A:Gene: pdhR
 C:Superfamily: pyruvate dehydrogenase complex repressor

Query Match 2.0%; Score 8; DB 2; Length 254;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 REAIQRL 366
 |||||
 DB 52 REAIQRL 59

RESULT 13

AF0521
 Pyruvate dehydrogenase complex repressor [imported] - Salmonella enterica subsp. enter
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: This species has also been called Salmonella typhi
 C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AF0521
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
 th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Parre
 , S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, P.
 A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica ser
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AF0521
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-254 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD01310.1; PID:gl6501438; GSPDB:GN00176
 C:Genetics:
 A:Gene: STY0174
 C:Superfamily: pyruvate dehydrogenase complex repressor

Query Match 2.0%; Score 8; DB 2; Length 254;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 REAIQRL 366
 |||||
 DB 52 REAIQRL 59

RESULT 14

E82079
 Pyruvate dehydrogenase complex repressor VC2415 [imported] - Vibrio cholerae (strain E82079)
 C:Species: Vibrio cholerae
 C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 19-Jul-2002
 C:Accession: E82079

R.Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardonson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B. I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A>Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
 A;Reference number: AB2035; MUID:20406833; PMID:10952301
 A;Accession: E82079
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-256 <HEI>
 A;Cross-references: GB:AE004311; GB:AE003952; NID:g9656983; PIDN:AAF95558.1; GSPDB:GN001
 A;Experimental source: serogroup O1; strain N16961; biotype El Tor
 C;Genetics:
 A;Gene: VC2415
 A;Map position: 1
 C;Superfamily: pyruvate dehydrogenase complex repressor

Query Match 2.0%; Score 8; DB 2; Length 256;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 REAIOQLE 366
 |||||
 Db 52 REAIOQLE 59

RESULT 15
 B23034
 streptomycin resistance protein - *Klebsiella* sp. transposon Tn5
 C;Species: *Klebsiella* sp.
 C;Date: 20-Aug-1987 #sequence_revision 20-Aug-1987 #text_change 19-Feb-2000
 C;Accession: B23034
 R;Mazodier, P.; Cossart, P.; Gitraud, E.; Gasser, F.
 Nucleic Acids Res. 13 195-205, 1985
 A>Title: Completion of the nucleotide sequence of the central region of Tn5 confirms the
 A;Reference number: A93559; MUID:85215465; PMID:3889831
 A;Accession: B23034
 A;Molecule type: DNA
 A;Residues: 1-266 <MAZ>
 A;Superfamily: streptomycin 6-kinase
 C;Keywords: antibiotic resistance

Query Match 2.0%; Score 8; DB 2; Length 266;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 APAASVAR 111
 |||||
 Db 132 APAASVAR 139

RESULT 16
 T03254
 probable carbonate dehydratase (EC 4.2.1.1), chloroplast - rice
 N;Alternate names: carbonic anhydrase
 C;Species: *Oryza sativa* (rice)
 C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 22-Jun-1999
 C;Accession: T03254
 R;Suzuki, S.; Burnell, J.N.
 submitted to the EMBL Data Library, April 1994
 A;Description: Rice carbonic anhydrase: purification, location, and sequence analysis of
 A;Reference number: Z14863
 A;Accession: T03254
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-273 <SUZ>
 A;Cross-references: EMBL:U08404; NID:g606816; PIDN:AAA86943.1; PID:g606817
 A;Experimental source: leaf
 C;Genetics:
 A;Genome: nuclear
 C;Function:
 A;Description: catalyzes the reversible hydration of carbon dioxide
 C;Superfamily: *Escherichia coli* carbonate dehydratase

C;Keywords: carbon-oxygen lyase; chloroplast; hydro-lyase

Query Match 2.0%; Score 8; DB 2; Length 273;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 AAPVAPAA 107
 |||||
 Db 64 AAPVAPAA 71

RESULT 17
 AE2006
 hypothetical protein alr1603 [imported] - *Nostoc* sp. (strain PCC 7120)
 C;Species: *Nostoc* sp. PCC 7120
 A;Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
 C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C;Accession: AE2006
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguc
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,
 DNA Res. 8, 205-213, 2001
 A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium A
 A;Reference number: AB1807; MUID:21595285; PMID:11759840
 A;Accession: AE2006
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-289 <KUR>
 A;Cross-references: GB:BA000019; PIDN:BA877969.1; PID:g17135423; GSPDB:GN00179
 A;Experimental source: strain PCC 7120
 C;Genetics:
 A;Gene: alr1603

Query Match 2.0%; Score 8; DB 2; Length 289;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 AAPAATVA 139
 |||||
 Db 56 AAPAATVA 63

RESULT 18
 T08106
 2'-hydroxyisoflavone reductase (EC 1.3.1.45) - European white birch
 C;Species: *Betula pendula* (European white birch)
 C;Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 26-May-2000
 R;Schmitz, N.; Frank, E.; Scheurer, S.; Hausteiner, D.; Vieths, S.
 submitted to the EMBL Data Library, February 1998
 A;Description: Characterization of a new birch pollen allergen, representing an IPRH p
 A;Reference number: Z16359
 A;Accession: T08106
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-300 <SCH>
 A;Cross-references: EMBL:AF047896; NID:g2935415; PIDN:AAC05116.1; PID:g2935416
 C;Superfamily: 2'-hydroxyisoflavone reductase
 C;Keywords: oxidoreductase

Query Match 2.0%; Score 8; DB 2; Length 300;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 TNFEIEAS 18
 |||||
 Db 276 TNFEIEAS 283

RESULT 19
 T34406
 hypothetical protein EGAP9.3 - *Caenorhabditis elegans*
 C;Species: *Caenorhabditis elegans*
 C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T34406
R:Geisel, C.
A:Submitted to the EMBL Data Library, November 1996
A:Description: The sequence of *C. elegans* cosmid EGAP9.
A:Reference number: Z21519
A:Accession: T34406
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-339 <GEI>
A:Cross-references: EMBL:U80026; PIDN:RAC25845.1; GSPDB:GN00023; CESP:EGAP9.3
A:Experimental source: strain Bristol N2; clone EGAP9
C:Genetics:
A:Gene: CESP:EGAP9.3
A:Map position: 5
A:Introns: 40/3; 78/3; 117/2; 152/1; 208/1; 238/3; 267/2; 309/2

Query Match 2.0%; Score 8; DB 2; Length 339;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 RAIDLYLS 200
|||||
Db 143 RAIDLYLS 150

RESULT 20
S59499
cellulase eg11 - smut fungus (*Ustilago maydis*)
C:Species: *Ustilago maydis* (corn smut)
C:Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 28-May-1999
C:Accession: S59499
R:Schauwecker, F.; Wanner, G.; Kahmann, R.
Biol. Chem. Hoppe-Seyler 376, 617-625, 1995
A:Title: Filament-specific expression of a cellulase gene in the dimorphic fungus *Ustilago maydis*
A:Reference number: S59499; MUID:96145728; PMID:8590631
A:Accession: S59499
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-393 <SCH>
A:Cross-references: GB:S81598; NID:gl478378; PIDN:AAB36147.1; PID:gl478379

Query Match 2.0%; Score 8; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 PSQGSNPG 266
|||||
Db 362 PSQGSNPG 369

RESULT 21
A84327
photolyase/cryptochrome [imported] - *Halobacterium* sp. NRC-1
C:Species: *Halobacterium* sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: A84327
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablonski, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of *Halobacterium* species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: A84327
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-462 <STO>
A:Cross-references: GB:AE004437; NID:gl0581208; PIDN:AAG19981.1; GSPDB:GN00138
C:Genetics:
A:Gene: phr1
C:Superfamily: deoxyribodipyrimidine photo-lyase

Query Match 2.0%; Score 8; DB 2; Length 462;

Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 181 VRALRAAY 188
|||||
Db 53 VRALRAAY 60

RESULT 22
S74633
high affinity sulfate transporter - *Synechocystis* sp. (strain PCC 6803)
A:Alternate names: hypothetical protein slr1776
C:Species: *Synechocystis* sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C:Accession: S74633
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasu
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74633
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-566 <KAN>
A:Cross-references: EMBL:D90900; GB:AB001339; NID:gl651768; PIDN:BAA16785.1; PID:gl651
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Start codon: GTG
C:Superfamily: integral membrane protein HP0228

Query Match 2.0%; Score 8; DB 2; Length 566;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 GVPSSGGSN 264
|||||
Db 155 GVPSSGGSN 162

RESULT 23
D70697
probable membraneprotein - *Mycobacterium tuberculosis* (strain H37RV)
C:Species: *Mycobacterium tuberculosis*
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: D70697
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, J.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, R.; Andreaman, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
C:Accession: D70697
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-643 <COL>
A:Cross-references: GB:Z80343; GB:AL123456; NID:gl3261648; PIDN:CAB02471.1; PID:e267405
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv3792

Query Match 2.0%; Score 8; DB 2; Length 643;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 LQQLAAAV 348
|||||
Db 29 LQQLAAAV 36

RESULT 24

AI3404
transposase BME11223 [imported] - Brucella melitensis (strain 16M)
C/Species: Brucella melitensis
C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C/Accession: AI3404
R:/DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, M.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, J.; Mouton, M.; U.S.A. 99, 443-448, 2002
A/Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A/Reference number: AD352; PMID:11756688
A/Accession: AI3404
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-653 <KUR>
A/Cross-references: GB:AE008917; PIDN:AAL52404.1; PID:gl7983205; GSPDB:GN00190
A/Experimental source: strain 16M
C/Genetics:
A/Gene: BME11223
A/Map position: I

Query Match 2.0%; Score 8; DB 2; Length 653;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 273 SGALDALR 280
|||||
Db 255 SGALDALR 262

RESULT 25
D87270
general secretion pathway protein D [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C/Accession: D87270
R:/Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J. J.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonin, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647
A/Accession: D87270
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-687 <STO>
A/Cross-references: GB:AE005673; NID:gl3421292; PIDN:AAK22160.1; GSPDB:GN00148
C/Genetics:
A/Gene: CC0173

Query Match 2.0%; Score 8; DB 2; Length 687;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 GASTATTA 90
|||||
Db 307 GASTATTA 314

RESULT 26
DJBEV1
DNA-directed DNA polymerase (EC 2.7.7.7) - human herpesvirus 1
C/Species: human herpesvirus 1
A/Note: host Homo sapiens (man)
C/Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 28-May-1999
C/Accession: A00715
R:/Quinn, J.P.; McGeoch, D.J. Nucleic Acids Res. 13, 8143-8163, 1985
A/Title: DNA sequence of the region in the genome of herpes simplex virus type 1 containing the unique long (UL) gene
A/Reference number: A93601; MUID:86067223; PMID:2999714
A/Accession: A00715
A/Molecule type: DNA
A/Residues: 1-1235 <QUI>
A/Cross-references: GB:X03181; GB:M12356; NID:G59862; PIDN:CAA26941.1; PID:G59864

RESULT 29

C84025
 hypothetical protein BH3003 [imported] - Bacillus halodurans (strain C-125)
 C:Species: Bacillus halodurans
 C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: C84025
 R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A:Reference number: A83650; MUID:20512582; PMID:11058132
 A:Accession: C84025
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-58 <STO>
 A:Cross-references: GB:AP001517; GB:BA000004; NID:G10175500; PIDN:BAB06722.1; GSPDB:GN00
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH3003

Query Match 1.7%; Score 7; DB 2; Length 58;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 357 EREAIQ 363
 |||||
 Db 10 EREAIQ 16

RESULT 30

AG1930
 hypothetical protein asl0994 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C:Accession: AG1930
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nukazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AG1930
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-87 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BAB72951.1; PID:G17130340; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: asl0994

Query Match 1.7%; Score 7; DB 2; Length 87;
 Best Local Similarity 100.0%; Pred. No. 43;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 325 LRLVNES 331
 |||||
 Db 5 LRLVNES 11

RESULT 31

B89979
 conserved hypothetical protein SA1723 [imported] - Staphylococcus aureus (strain N315)
 C:Species: Staphylococcus aureus
 C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 C:Accession: B89979
 R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogu
 ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hitamatsu, K.
 Lancet 357, 1225-1240, 2001
 A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
 A:Reference number: A89758; MUID:21311952; PMID:11418146
 A:Accession: B89979
 A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-100 <KUR>
 A:Cross-references: GB:BA000018; PID:G13701700; PIDN:BAB42993.1; GSPDB:GN00149
 A:Experimental source: strain N315
 C:Genetics:

A:Gene: SA1723

C:Superfamily: Bacillus subtilis hypothetical protein yerc

Query Match 1.7%; Score 7; DB 2; Length 100;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 SGASTAT 88
 |||||

Db 66 SGASTAT 72
 |||||

RESULT 32

AG0699
 probable pathogenicity island protein sseA [imported] - Salmonella enterica subsp. ent
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: this species has also been called Salmonella typhi
 C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AG0699
 R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farra
 . S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica ser
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AG0699
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-108 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD01968.1; PID:G16502810; GSPDB:GN00176
 C:Genetics:
 A:Gene: sseA

Query Match 1.7%; Score 7; DB 2; Length 108;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 252 NLFPOGV 258
 |||||

Db 59 NLFPOGV 65
 |||||

RESULT 33

G91250
 hypothetical protein ECs4975 [imported] - Escherichia coli (strain O157:H7, substrain
 C:Species: Escherichia coli
 C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
 C:Accession: G91250
 R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and ge
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: G91250
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-128 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BAB38398.1; PID:G13364451; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: ECs4975

Query Match 1.7%; Score 7; DB 2; Length 128;
 Best Local Similarity 100.0%; Pred. No. 61;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 TVAATDD 143

C;Keywords: blocked amino end; cytosol; NAD; oxidoreductase

Query Match 1.7%; Score 7; DB 2; Length 144;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 227 SPAQPAV 233
|
Db 2 SPAQPAV 8
|

RESULT 37
AB2443
hypothetical protein all5098 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Date: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Note: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AB2443
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigun,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata,
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium A
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AB2443
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-144 <KUR>
A;Cross-references: GB:BAO00019; PIDN:BAB76797.1; PID:g17134236; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all5098

Query Match 1.7%; Score 7; DB 2; Length 144;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 118 VATAETA 124
|
Db 24 VATAETA 30
|

RESULT 38
D86749
riboflavin synthase (EC 2.5.1.9) beta chain [imported] - Lactococcus lactis subsp. lac
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C;Accession: D86749
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehr
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: D86749
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-154 <STO>
A;Cross-references: GB:AE005176; PID:g12723940; PIDN:AAK05094.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: ribH
C;Superfamily: riboflavin synthase beta chain
C;Keywords: transferase

Query Match 1.7%; Score 7; DB 2; Length 154;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 273 SGALDAL 279
|
Db 32 SGALDAL 38
|

RESULT 39
C90281

hypothetical protein SS01269 [imported] - Sulfolobus solfataricus
 C:Species: Sulfolobus solfataricus
 C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
 C:Accession: C90281
 R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.
 arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
 submitted to GenBank, April 2001
 A:Description: Sulfolobus solfataricus complete genome.
 A:Reference number: A99139
 A:Accession: C90281
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-163 <KUR>
 A:Cross-references: GB:AE006641; NID:g13814464; PIDN:AAK41506.1; GSPDB:GN00155
 C:Genetics:
 A:Gene: SS01269

Query Match 1.7%; Score 7; DB 2; Length 163;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 DVKRIIE 31
 |||||
 DB 101 DVKRIIE 107

RESULT 40
 F75363
 hypothetical protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 26-May-2000
 C:Accession: F75363
 R;White, O.; Eisen, J.A.; Heideberg, J.P.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: F75363
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-176 <WHI>
 A:Cross-references: GB:AE002012; GB:AE000513; NID:g6459473; PIDN:AAF11269.1; PID:g645948
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR1708
 A:Map position: 1
 C:Superfamily: Deinococcus radiodurans hypothetical protein DR1708

Query Match 1.7%; Score 7; DB 2; Length 176;
 Best Local Similarity 100.0%; Pred. No. 82;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 QAAPAT 137
 |||||
 DB 27 QAAPAT 33

RESULT 41
 E95274
 hypothetical protein Sma0191 [imported] - Sinorhizobium meliloti (strain 1021) magapla
 C:Species: Sinorhizobium meliloti
 C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
 C:Accession: E95274
 R;Barnett, M.J.; Jones, R.F.; Fisher, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bow
 ; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
 Proc. Natl. Acad. Sci. U.S.A. 98, 9893-9898, 2001
 A>Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilo
 A:Reference number: A95262; MUID:21396509; PMID:11481432
 A:Accession: E95274
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-176 <KUR>
 A:Cross-references: GB:AE006469; PIDN:AAK64759.1; PID:g14523165; GSPDB:GN00165
 A:Experimental source: strain 1021, megaplasmid pSymA
 R;Galibert, F.; Finan, T.M.; Long, S.R.; Fuhler, A.; Abola, P.; Ampe, F.; Barloy-Huble
 pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kias, E.; Komp, C.; Lelau
 hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
 A>Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: Sma0191
 A:Genome: plasmid

Query Match 1.7%; Score 7; DB 2; Length 176;
 Best Local Similarity 100.0%; Pred. No. 82;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 AAPAATV 138
 |||||
 DB 18 AAPAATV 24

RESULT 42
 F71304
 probable ribosome recycling factor - syphilis spirochete
 C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
 C>Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 26-Aug-1999
 C:Accession: F71304
 R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gw
 rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; Mc
 they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
 Science 281, 375-388, 1998
 A>Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
 A:Reference number: A71250; MUID:98332770; PMID:9665876
 A:Accession: F71304
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-183 <COL>
 A:Cross-references: GB:AE001235; GB:AE000520; NID:g3322893; PIDN:AAC65577.1; PID:g3322
 A:Experimental source: strain Nicholas
 C:Genetics:
 A:Gene: TP0604
 C:Superfamily: ribosome releasing factor

Query Match 1.7%; Score 7; DB 2; Length 183;
 Best Local Similarity 100.0%; Pred. No. 85;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 DASVADV 26
 |||||
 DB 163 DASVADV 169

RESULT 43
 IS1347
 MHC class I - Atlantic salmon (fragment)
 C:Species: Salmo salar (Atlantic salmon)
 C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 03-Dec-1999
 C:Accession: IS1347
 R;Grimholt, U.; Hordvik, I.; Fosse, V.M.; Olsaker, I.; Endresen, C.; Lie, O.
 Immunogenetics 37, 469-473, 1993
 A>Title: Molecular cloning of major histocompatibility complex class I cDNAs in Atlan
 A:Reference number: IS1347; MUID:93170896; PMID:8436423
 A:Accession: IS1347
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-188 <GRI>
 A:Cross-references: GB:L07605; NID:g213820; PIDN:AAA49602.1; PID:g213821
 C:Genetics:
 A:Gene: MHC-I

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 1.7%; Score 7; DB 2; Length 188;
 Best Local Similarity 100.0%; Pred. No. 87;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 352 LVTVPPE 358
 |||||
 Db 106 LVTVPPE 112

RESULT 44

H84364
 membrane protein [imported] - Halobacterium sp. NRC-1
 C:Species: Halobacterium sp. NRC-1
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: H84364
 R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablonski, K.H.; Alam, M.; Freitas, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
 A:Title: Genome sequence of Halobacterium species NRC-1.
 A:Reference number: A84160; MUID:20504483; PMID:11016950
 A:Accession: H84364
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-218 <STO>
 A:Cross-references: GB:AE004437; NID:gl0581563; PIDN:AAG20284.1; GSPDB:GN00138
 C:Genetics: hmp

Query Match 1.7%; Score 7; DB 2; Length 218;
 Best Local Similarity 100.0%; Pred. No. 99;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 211 VARAPAA 217
 |||||
 Db 210 VARAPAA 216

RESULT 45

T29210
 hypothetical protein T20F5.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T29210
 R:Du, Z.; Le, T.T.
 submitted to the EMBL Data Library, November 1996
 A:Description: The sequence of C. elegans coamid T20F5.
 A:Reference number: 220588
 A:Accession: T29210
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-226 <DUZ>
 A:Cross-references: EMBL:U80442; PIDN:AAB37667.1; GSPDB:GN00019; CESP:T20F5.5
 A:Experimental source: strain Bristol N2; clone T20F5
 C:Genetics:
 A:Gene: CESP:T20F5.5
 A:Map position: 1
 A:Introns: 14/3

Query Match 1.7%; Score 7; DB 2; Length 226;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 99 PAAPVAP 105
 |||||
 Db 125 PAAPVAP 131

RESULT 46

T45388

hypothetical protein [imported] - Mycobacterium leprae

C:Species: Mycobacterium leprae
 C:Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 31-Jan-2000
 C:Accession: T45388
 R:Cole, S.T.; Flesselles, B.; Honore, N.
 submitted to the EMBL Data Library, August 1997
 A:Reference number: Z22966
 A:Accession: T45388
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-227 <COL>
 A:Cross-references: EMBL:Z98756; PIDN:CAB11458.1
 A:Experimental source: cosmid B2492
 C:Genetics:
 A:Note: MLCB2492.26

Query Match 1.7%; Score 7; DB 2; Length 227;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 211 VARAPAA 217
 |||||
 Db 212 VARAPAA 218

RESULT 47

S46965
 microfilarial sheath protein - nematode (Brugia pahangi)
 C:Species: Brugia pahangi
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 19-May-2000
 C:Accession: S46965
 R:Hirzmann, J.; Schnauffer, A.; Hintz, M.; Conratha, F.J.; Stirm, S.; Zahner, H.; Hobo
 submitted to the EMBL Data Library, July 1994
 A:Description: Brugia spp. and Litomosoides carinii: characterization of the shp2 gene
 A:Reference number: S46964
 A:Accession: S46965
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-228 <HIR>
 A:Cross-references: EMBL:Z35445; NID:gs16354; PID:gs16355
 C:Genetics:
 A:Introns: 26/1
 C:Superfamily: hydroxyproline-rich glycoprotein

Query Match 1.7%; Score 7; DB 2; Length 228;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 99 PAAPVAP 105
 |||||
 Db 81 PAAPVAP 87

RESULT 48

E86338
 protein F5M15.16 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: E86338
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K
 anson, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziah
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: E86338
 A:Status: preliminary
 A:Molecule type: DNA

```

A:Residues: 1-231 <STO>
A:Cross-references: GB:AE005172; NID:98778602; PIDN:AAF79610.1; GSPDB:GN00141
C:Genetics:
A:Gene: F5M15.16
A:Map position: 1

Query Match      1.7%; Score 7; DB 2; Length 231;
Best Local Similarity 100.0%; Pred. No. 1.e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 390 ANYLLDH 396
    |||||
Db 16 ANYLLDH 22

RESULT 49
A:Accession: AI3283
transcription regulator, gntr family BMEI0254 [imported] - Brucella melitensis (strain 1
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AI3283
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letese
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A:Reference number: AD3252; PMID:11756688
A:Accession: AI3283
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-231 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAL52417.1; PID:G17983220; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI0254
A:Map position: 1

Query Match      1.7%; Score 7; DB 2; Length 231;
Best Local Similarity 100.0%; Pred. No. 1.e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 REAIQRL 365
    |||||
Db 70 REAIQRL 76

RESULT 50
A:Accession: AF3406
hypothetical exported proline-rich protein BMEI1236 [imported] - Brucella melitensis (st
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AF3406
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letese
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A:Reference number: AD3252; PMID:11756688
A:Accession: AF3406
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-232 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAL52417.1; PID:G17983220; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI1236
A:Map position: 1

Query Match      1.7%; Score 7; DB 2; Length 232;
Best Local Similarity 100.0%; Pred. No. 1.e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 PAAPVAP 105
    |||||
Db 75 PAAPVAP 81

```

```

RESULT 51
H83559
hypothetical protein PA0674 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: H83559
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Li
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pat
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: H83559
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-238 <STO>
A:Cross-references: GB:AE004503; GB:AE004091; NID:G9946553; PIDN:AAG04063.1; GSPDB:GNO
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0674

Query Match      1.7%; Score 7; DB 2; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 PATLAQP 99
    |||||
Db 207 PATLAQP 213

RESULT 52
E70822
hypothetical protein Rv0730 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: E70822
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genom
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: E70822
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-242 <COL>
A:Cross-references: GB:AL021958; GB:AL123456; NID:G3261536; PIDN:CAA17497.1; PID:G2911
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv0730

Query Match      1.7%; Score 7; DB 2; Length 242;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 VARAPAA 217
    |||||
Db 226 VARAPAA 232

RESULT 53
AD0329
conserved hypothetical protein YPO2700 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C:Accession: AD0329
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell
Nature 413, 523-527, 2001

```

A>Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A;Reference number: AB0001; MUID:21470413; PMID:11586360
 A;Accession: AD0329
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-245 <KUR>
 A;Cross-references: GB:AL590842; PIDN:CAC92939.1; PID:gl5980678; GSPDB:GN00175
 C;Genetics:
 A;Gene: YPO2700
 C;Superfamily: Bacillus subtilis lactam utilization protein ycsF

Query Match 1.7%; Score 7; DB 2; Length 245;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 286 QALLQLV 292
 Db 17 QALLQLV 23

RESULT 54
 AC0830

Conserved hypothetical protein STY2835 [imported] - Salmonella enterica subsp. enterica
 C;Species: Salmonella enterica subsp. enterica serovar Typhi
 A;Note: this species has also been called Salmonella typhi
 C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 A;Accession: AC0830
 R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moule, S.; O'Garra, P.
 Nature 413, 848-852, 2001
 A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
 A;Reference number: AB0502; MUID:21534947; PMID:11677608
 A;Accession: AC0830

A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-245 <PAR>

A;Cross-references: GB:AL513382; PIDN:CAD02791.1; PID:gl6503801; GSPDB:GN00176
 C;Genetics:
 A;Gene: STY2835

C;Superfamily: hypothetical protein HI0423

Query Match 1.7%; Score 7; DB 2; Length 245;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 VADVKKI 29
 Db 42 VADVKKI 48

RESULT 55
 T50874

hypothetical cytochrome [imported] - Rubrivivax gelatinosus
 C;Species: Rubrivivax gelatinosus
 C;Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
 A;Accession: T50874
 R;Nagashima, K.V.; Igarashi, N.; Harada, J.; Nagashima, S.; Matsuura, K.; Shimada, K.
 submitted to the EMBL Data Library, November 1999
 A;Description: Determination of Nucleotide Sequences of Rubrivivax gelatinosus Photosynt
 A;Reference number: Z25270
 A;Accession: T50874
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-247 <NAG>
 A;Cross-references: EMBL:AB034704; PIDN:BAA94027.1
 C;Genetics:
 A;Note: ORF247

Query Match 1.7%; Score 7; DB 2; Length 247;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 100 AAPVAPA 106
 Db 48 AAPVAPA 54

RESULT 56
 JCS237

osmotin-like protein precursor - tomato
 C;Species: Lycopersicon esculentum (tomato)
 C;Date: 13-Mar-1997 #sequence_revision 13-Mar-1997 #text_change 21-Jul-2000
 A;Accession: JCS237
 R;Chen, R.; Wang, F.; Smith, A.G.
 Gene 179, 301-302, 1996

A;Title: A flower-specific gene encoding an osmotin-like protein from Lycopersicon esc
 A;Reference number: JCS237; MUID:97128324; PMID:8972917

A;Contents: flower
 A;Accession: JCS237
 A;Status: preliminary; nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Residues: 1-252 <CHE>

A;Cross-references: GB:L76632; NID:gl220536; PIDN:AAB41124.1; PID:gl220537
 C;Comment: This protein is rich in cysteine content and play a part of the flowers def.
 C;Superfamily: thaumatin I
 F;1-23/Domain: signal sequence #status predicted <SIG>
 F;24-252/Product: osmotin-like protein #status predicted <MAT>

Query Match 1.7%; Score 7; DB 2; Length 252;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 92 APATLAQ 98
 Db 113 APATLAQ 119

RESULT 57
 T29527

hypothetical protein F27C1.2 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 A;Accession: T29527

R;Wu, X.; Le, T.T.
 submitted to the EMBL Data Library, November 1996
 A;Description: The sequence of C. elegans cosmid F27C1.
 A;Reference number: Z20636

A;Accession: T29527
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA

A;Residues: 1-252 <WUX>
 A;Cross-references: EMBL:U80441; PIDN:AAB37655.1; GSPDB:GN00019; CESP:F27C1.2
 A;Experimental source: strain Bristol N2; clone F27C1

C;Genetics:

A;Gene: CESP:F27C1.2
 A;Map position: 1
 A;Introns: 23/3; 100/1; 142/3; 216/1

Query Match 1.7%; Score 7; DB 2; Length 252;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 107 ASVARTP 113
 Db 172 ASVARTP 178

RESULT 58
 C82868

plasmid replication protein XFa0060 [imported] - Xylella fastidiosa (strain 9a5c)
 C;Species: Xylella fastidiosa
 C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 24-May-2001
 C;Accession: C82868

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
 A:Reference number: A82515; MUID:20365717; PMID:10910347
 A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: C82868
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-254 <SIM>
 A:Cross-references: GB:AE003851; NID:G9112238; PIDN:AAF95628.1; GSPDB:GN00130; XFSC:XFA0
 A:Experimental source: strain 945c
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
 ae-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Tshahko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
 A:Reference number: A59328
 A:Content: annotation
 C:Genetics:
 A:Gene: XFA0060
 A:Genome: plasmid
 A:Note: plasmid pXF5.1
 C:Superfamily: regulatory protein spo0J

Query Match 1.7%; Score 7; DB 2; Length 254;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 275 ALDALRQ 281
 |||||
 Db 192 ALDALRQ 198
 |||||

RESULT 59
 T25073
 hypothetical protein T21C9.9 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T25073
 R:McMurray, A.
 submitted to the EMBL Data Library, May 1996
 A:Reference number: Z19977
 A:Accession: T25073
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-254 <WIL>
 A:Cross-references: EMBL:Z73098; PIDN:CAA97336.1; GSPDB:GN00023; CESP:T21C9.9
 A:Experimental source: clone T21C9
 C:Genetics:
 A:Gene: CESP:T21C9.9
 A:Map position: 5
 A:Introns: 22/2; 78/3

Query Match 1.7%; Score 7; DB 2; Length 254;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 QPAAPVA 104
 |||||
 Db 133 QPAAPVA 139
 |||||

RESULT 60
 S76756
 1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino]imidazole-4-carboxamide
 C:Species: *Synechocystis* sp.

A:Variety: PCC 6803
 C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 03-Jun-2000
 A:Accession: S76756
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N
 o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasu
 DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocyst*
 s.
 A:Reference number: S74322; MUID:97061201; PMID:8905231
 A:Accession: S76756
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-256 <KAN>
 A:Cross-references: EMBL:D90916; GB:AB001339; NID:G1653715; PIDN:BA18668.1; PID:G1653
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C:Superfamily: N-(5'-phospho-D-riboseylformimino)-5-amino-1-(5'-phosphoribosyl)-4-imid
 C:Keywords: intramolecular oxidoreductase; isomerase

Query Match 1.7%; Score 7; DB 2; Length 256;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 276 LDALRQL 282
 |||||
 Db 179 LDALRQL 185
 |||||

RESULT 61
 B32393
 T-cell antigen 4-1BB precursor - mouse
 C:Species: *Mus musculus* (house mouse)
 C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 21-Jul-2000
 C:Accession: B32393; I48879
 R:Kwon, B.S.; Weisman, S.M.
 Proc. Natl. Acad. Sci. U.S.A. 86, 1963-1967, 1989
 A:Title: cDNA sequence of two inducible T-cell genes.
 A:Reference number: A32393; MUID:89184547; PMID:2784565
 A:Accession: B32393
 A:Molecule type: mRNA
 A:Residues: 1-256 <KNO>
 A:Cross-references: GB:J04492; NID:G201121; PIDN:AAA40167.1; PID:G201122
 R:Kwon, B.S.; Kozak, C.A.; Kim, K.K.; Pickard, R.T.
 J. Immunol. 152, 2256-2262, 1994
 A:Title: Genomic Organization and Chromosomal Localization of the T-Cell Antigen 4-1B
 A:Reference number: I48879; MUID:94179805; PMID:8133039
 A:Accession: I48879
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-256 <RES>
 A:Cross-references: EMBL:U02567; NID:G1117783; PIDN:AAA93113.1; PID:G409178
 C:Genetics:
 A:Introns: 35/1; 69/1; 115/1; 138/2; 156/3; 181/1; 225/1
 C:Superfamily: CD27 antigen; NGF receptor repeat homology
 C:Keywords: transmembrane protein
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-256/Product: 4-1BB protein #status predicted <MAT>

Query Match 1.7%; Score 7; DB 2; Length 256;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 PEGPGGG 338
 |||||
 Db 175 PEGPGGG 181
 |||||

RESULT 62
 T46548
 probable transcription activator [imported] - *Streptomyces violaceoruber*
 C:Species: *Streptomyces violaceoruber*
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 11-May-2000
 C:Accession: T46548
 R:Ichinose, K.; Bedford, D.J.; Tornus, D.; Bechthold, A.; Bibb, M.J.; Revill, W.P.; F

Chem. Biol. 5, 647-659, 1998

A:Title: The granaticin biosynthetic gene cluster of Streptomyces violaceoruber Tu22: see
A:Reference number: 223045; MUID:99051446; PMID:9831526
A:Accession: T46548
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-257 <ICH>
A:Cross-references: EMBL:AJ011500; PIDN:CAA09664.1
A:Experimental source: strain Tu22
C:Genetics:
A:Note: gra-orf37

Query Match 1.7%; Score 7; DB 2; Length 257;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 272 GSGALDA 278
Db 224 GSGALDA 230

RESULT 63

A75618
Probable chromosome partitioning ATPase, ParA family - Deinococcus radiodurans (strain R
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
A:Accession: A75618
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
S.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: A75618
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-258 <WHI>
A:Cross-references: GB:AE001826; NID:G6460827; PIDN:AAF12609.1; PID:G6460905; TIGR:DRB00
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRB0001
A:Map position: megaplasmid
A:Genome: plasmid
A:Note: plasmid MPI
C:Superfamily: regulatory protein spo0J

Query Match 1.7%; Score 7; DB 2; Length 258;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 341 LGQLAAA 347
Db 127 LGQLAAA 133

RESULT 64

G83245
Probable ATP-binding component of ABC transporter PA3212 [imported] - Pseudomonas aerugi
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
A:Accession: G83245
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: G83245
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-264 <STO>
A:Cross-references: GB:AE004744; GB:AE004091; NID:G9949317; PIDN:ARG06600.1; GSPDB:GN001
A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA3212

Query Match 1.7%; Score 7; DB 2; Length 264;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 138 VAATDDA 144
Db 234 VAATDDA 240

RESULT 65

T10644
Hypothetical protein T13K14.180 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 15-Oct-1999
A:Accession: T10644
R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lenc
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z16991
A:Accession: T10644
A:Molecule type: DNA
A:Residues: 1-266 <BEV>
A:Cross-references: EMBL:AL080282; GSPDB:GN00062; ATSP:T13K14.180
A:Experimental source: cultivar Columbia; BAC clone T13K14
C:Genetics:
A:Gene: ATSP:T13K14.180
A:Map position: 4
A:Introns: 47/3

Query Match 1.7%; Score 7; DB 2; Length 266;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 91 KAPATLA 97
Db 22 KAPATLA 28

RESULT 66

S72844
beta-aspartate methyltransferase pimT - Mycobacterium leprae
N:Alternate names: B2126 C1 165 protein
C:Species: Mycobacterium leprae
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 23-Mar-2001
A:Accession: S72844
R:Smith, D.R.; Robison, K.
submitted to the EMBL Data Library, November 1993
A:Description: Mycobacterium leprae cosmid B2126.
A:Reference number: S72585
A:Accession: S72844
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-280 <SMI>
A:Cross-references: EMBL:U00017; NID:G466994; PIDN:AAA17184.1; PID:G466999
C:Genetics:
A:Gene: pimT
A:Start codon: GTG

Query Match 1.7%; Score 7; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 270 GAGSGAL 276
Db 107 GAGSGAL 113

RESULT 67

A70513
Hypothetical protein RV2118c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis

RESULT 71
 AG3297
 biopolymer transport exbB protein [imported] - *Brucella melitensis* (strain 16M)
 C:Species: *Brucella melitensis*
 C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 17-May-2002
 C:Accession: AG3297
 R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova
 .; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Lefe
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melit*
 A:Reference number: AB3252; PMID:11756688
 A:Accession: AG3297
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-290 <KUR>
 A:Cross-references: GB:AE008917; PIDN:AAL51546.1; PID:gl7982264; GSPDB:GN00190

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMEI0365

A:Map position: 1

C:Superfamily: biopolymer transport protein

Query Match 1.7%; Score 7; DB 2; Length 290;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 AAPVAPA 106

|||||

Db 28 AAPVAPA 34

RESULT 72

TI8584

F15C11.2 - Caenorhabditis elegans (fragment)

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: TI8584

R:Barlow, K.

submitted to the EMBL Data Library, August 1997

A:Reference number: Z18990

A:Accession: TI8584

A>Status: preliminary; translated from GB/EMBL/DBPJ

A:Molecule type: DNA

A:Residues: 1-292 <WIL>

A:Cross-references: EMBL:Z99266; PIDN:CAB16464.1

A:Experimental source: Clone VF15C11L

C:Genetics:

A:Introns: 31/3; 67/1; 169/3; 212/3

Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 292;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 GKILKDE 56

|||||

Db 54 GKILKDE 60

RESULT 73

C96961

cell division protein (ftsX) [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001

C:Accession: C96961

R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: C96961

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-301 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK78478.1; PID:gl5023359; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC0498

Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 301;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 AATVAAT 141

|||||

Db 28 AATVAAT 34

RESULT 74

F81318

probable galactosyltransferase Cj1139c [imported] - Campylobacter jejuni (strain NCTC

C:Species: Campylobacter jejuni

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 06-Jan-2003

C:Accession: F81318

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Baeham, D.; Chill

C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barr

Nature 403, 665-668, 2000

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals h

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: F81318

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-303 <PAR>

A:Cross-references: GB:AL139077; GB:AL111168; NID:96968444; PIDN:CAB73394.1; PID:96968

A:Experimental source: serotype O2, strain NCTC 11168

C:Genetics:

A:Gene: Cj1139c

C:Superfamily: Neisseria meningitidis glycosyl transferase A

Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 303;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 ILDMGGG 173

|||||

Db 106 ILDMGGG 112

RESULT 75

F87629

hypothetical protein CC3072 [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: F87629

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kol

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: F87629

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-303 <STO>

A:Cross-references: GB:AE005673; NID:gl3424722; PIDN:AAK25034.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC3072

Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 303;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 PVAPAAAS 108

|||||

Db 59 PVAPAAAS 65

Search completed: December 17, 2003, 06:35:56

Job time : 24.3376 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 06:21:26 ; Search time 11.5265 Seconds
(without alignments)
1652.347 Million cell updates/sec

Title: US-09-805-550-2
Perfect score: 405
Sequence: 1 MKNLVKTLKGTNFEIASEPD.....BELTANYLLDHGHEFDQDQ 405

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	2.2	240	1 VANR_VIBAN	P74946 vibrio angu
2	9	2.2	264	1 RSPA_STRPU	P46771 strongyloce
3	9	2.2	674	1 PTGA_CORGL	Q45298 corynebacte
4	8	2.0	172	1 RL10_RHIO	Q98n68 rhizobium l
5	8	2.0	254	1 PDHR_ECOLI	P06957 escherichia
6	8	2.0	254	1 PDHR_SALTY	Q914b9 salmonella
7	8	2.0	266	1 STR_KLEPN	P13082 klebsiella
8	8	2.0	289	1 FRAH_ANASP	P46017 anabaena sp
9	8	2.0	346	1 MGBJ_HUMAN	O15480 homo sapien
10	8	2.0	393	1 GUNI_USTMA	P54424 ustilago ma
11	8	2.0	464	1 CAP_DICDI	P54654 dictyostell
12	8	2.0	804	1 BCB3_ACEXY	Q9w562 acetobacter
13	8	2.0	1235	1 DPOL_HSV11	P04293 herpes simp
14	7	1.7	124	1 SGP2_CHRV1	O52179 chromatium
15	7	1.7	126	1 CU24_ARADI	P80516 araneus dia
16	7	1.7	154	1 RISB_LACLA	Q9cgu6 lactococcus
17	7	1.7	156	1 RIB2_PHOLE	Q93e92 photobacter
18	7	1.7	183	1 RRP_TREPA	O83613 treponema p
19	7	1.7	211	1 KGUA_STRP3	Q8p001 streptococc
20	7	1.7	211	1 KGUA_STRP3	Q8p001 streptococc
21	7	1.7	252	1 OLPI_LYCES	Q99ym5 streptococc
22	7	1.7	256	1 HIS4_SYNY3	O41350 lycopersico
23	7	1.7	256	1 TNR9_MOUSE	P74561 synechocyst
24	7	1.7	280	1 LIGE_PSEPA	P20334 mus musculu
25	7	1.7	314	1 NARA_HUMAN	P27457 pseudomonas
26	7	1.7	314	1 NARA_HUMAN	Q93070 homo sapien
27	7	1.7	324	1 NARA_PANTR	Q95ne0 pan troglod
28	7	1.7	324	1 CAHC_HORVU	P40880 hordeum vul
29	7	1.7	335	1 G3P_HALN1	Q9hsx7 halobacteri
30	7	1.7	358	1 GBO_GEOCY	Q9x3v3 geodia cydo
31	7	1.7	360	1 POSI_BOVIN	O97552 bos taurus
32	7	1.7	364	1 INC2_ECOLI	P07673 escherichia
33	7	1.7	365	1 LA01_PONPY	P16211 pongo pygma
33	7	1.7	366	1 T2A1_ACICA	P24546 acinetobact

34	7	1.7	367	1	SGK2_MOUSE	Q9qz85 mus musculu
35	7	1.7	378	1	ADHP_RABIT	O46649 oryctolagus
36	7	1.7	386	1	SUCC_CHLTR	O84928 chlamydia t
37	7	1.7	395	1	EFTU_XYLPA	Q9p9g9 xylella fas
38	7	1.7	400	1	FXQ1_MOUSE	O70220 mus musculu
39	7	1.7	400	1	PROV_ECOLI	P14175 escherichia
40	7	1.7	400	1	PROV_SALTY	P17328 salmonella
41	7	1.7	427	1	SGK2_HUMAN	Q9hby8 homo sapien
42	7	1.7	448	1	YAP1_CHICK	P46936 gallus gall
43	7	1.7	465	1	ANT3_MOUSE	P32261 mus musculu
44	7	1.7	465	1	FXD3_MOUSE	O61060 mus musculu
45	7	1.7	484	1	COLB_MOUSE	Q9wm3 mus musculu
46	7	1.7	486	1	GALT_LACCA	O84904 lactobacill
47	7	1.7	498	1	YM22_CAEBL	Q21339 caenorhabdi
48	7	1.7	500	1	DHA1_MOUSE	P25459 mus musculu
49	7	1.7	500	1	DHA1_RAT	P51647 rattus norv
50	7	1.7	500	1	DHAC_RAT	P13601 rattus norv
51	7	1.7	507	1	IRX3_MOUSE	P81067 mus musculu
52	7	1.7	513	1	NQOD_PARDE	P29925 paracoccus
53	7	1.7	524	1	THS3_HALVO	Q9hha2 halobacteri
54	7	1.7	527	1	PTB_MOUSE	P17225 mus musculu
55	7	1.7	531	1	PTB_HUMAN	P26599 homo sapien
56	7	1.7	537	1	GAG_BAEVM	P03341 baboon endo
57	7	1.7	548	1	SYK_RHIO	Q98bc8 rhizobium l
58	7	1.7	551	1	SYK_BRUME	Q8yem8 bruceella su
59	7	1.7	551	1	SYK_BRUSU	P59225 bruceella su
60	7	1.7	555	1	PTB_RAT	Q00438 rattus norv
61	7	1.7	557	1	PTB_PIG	Q29099 sus scrofa
62	7	1.7	559	1	YHUJ_ECOLI	P37659 escherichia
63	7	1.7	597	1	Y745_ARATH	P59278 arabidopsis
64	7	1.7	629	1	SYR_PYRHO	O59147 pyrococcus
65	7	1.7	630	1	COG6_DROME	Q9v564 drosophila
66	7	1.7	641	1	PRIM_UREPA	Q9p266 ureaplasma
67	7	1.7	646	1	VP40_HSVB	P28336 equine herp
68	7	1.7	660	1	PKM1_COREF	Q8fui5 corynebacte
69	7	1.7	666	1	SCRT_DROME	P45843 drosophila
70	7	1.7	682	1	TUP1_KLULA	P56094 kluyveromyc
71	7	1.7	694	1	EFQ_BRUME	Q8yhp3 bruceella me
72	7	1.7	738	1	STII_YEAST	P23561 saccharomyc
73	7	1.7	750	1	ELS_CHICK	P07916 gallus gall
74	7	1.7	775	1	LVS4_EMENI	Q92412 emeritella
75	7	1.7	777	1	FRZE_MYXPO	P18769 myxococcus
76	7	1.7	801	1	VP34_SCHPO	P50520 schizosacch
77	7	1.7	812	1	P3K2_SOYBN	P42348 glycine max
78	7	1.7	814	1	P3K1_SOYBN	P42347 glycine max
79	7	1.7	814	1	P13K_ARATH	P42339 arabidopsis
80	7	1.7	816	1	P3K4_DICDI	P54676 dictyostell
81	7	1.7	857	1	APPI_SCHPO	Q9p7e8 schizosacch
82	7	1.7	870	1	BCA1_HUMAN	P56945 homo sapien
83	7	1.7	874	1	BCA1_MOUSE	Q61140 mus musculu
84	7	1.7	886	1	SYA_BARBA	P70865 bartonella
85	7	1.7	888	1	SYA_ZYMO	Q9rnn8 zymomonas m
86	7	1.7	917	1	AKH2_MAIZE	P49080 zea mays (m
87	7	1.7	920	1	AKH1_MAIZE	P49079 zea mays (m
88	7	1.7	964	1	MSH2_YEAST	P25847 saccharomyc
89	7	1.7	968	1	BCA1_RAT	Q63767 rattus norv
90	7	1.7	980	1	UBE4_CAEBL	Q09349 caenorhabdi
91	7	1.7	1004	1	PHC1_HUMAN	P78364 homo sapien
92	7	1.7	1012	1	PHC1_MOUSE	Q64028 mus musculu
93	7	1.7	1020	1	VP34_CANAL	Q92213 candida alb
94	7	1.7	1047	1	ANPB_BOVIN	P46197 bos taurus
95	7	1.7	1047	1	ANPB_HUMAN	P20594 homo sapien
96	7	1.7	1047	1	ANPB_RAT	P16067 rattus norv
97	7	1.7	1049	1	DSG1_HUMAN	Q02413 homo sapien
98	7	1.7	1055	1	POL_HV2ST	P20876 human immun
99	7	1.7	1061	1	RNE_ECOLI	P21513 escherichia
100	7	1.7	1070	1	Y355_HUMAN	O15063 homo sapien

ALIGNMENTS

RESULT 1

```

VANR VIBAN
ID VANR VIBAN STANDARD; PRT; 240 AA.
AC P74946;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transcriptional activator protein vanR.
GN VANR.
OS Vibrio anguillarum (Listonella anguillarum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Listonella.
OX NCBI_TaxID=55601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NB10 / Serotype O1;
RA MEDLINE=97284506; PubMed=9139920;
RX Milton D.L., Hardman A., Camara M., Chhabra S.R., Bycroft B.W.,
RA Stewart G.S.A.B., Williams P.;
RT "Quorum sensing in Vibrio anguillarum: characterization of the
RT vanI/vanR locus and identification of the autoinducer N-(3-
RT oxodecanoyl)-L-homoserine lactone.";
RL J. Bacteriol. 179:3004-3012(1997).
CC -|- SIMILARITY: PROBABLE TRANSCRIPTIONAL ACTIVATOR. BINDS TO AUTOINDUCER
CC MOLECULE ODHL.
CC -|- FUNCTION: BELONGS TO THE LUXR/UHPA FAMILY OF TRANSCRIPTIONAL
CC REGULATORS. AUTOINDUCER REGULATED SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U69677; AAC45213.1; -.
CC InterPro; IPR005143; Autoind bind.
CC InterPro; IPR000792; HTH LuxR.
CC Pfam; PF03472; Autoind_bind; 1.
CC Pfam; PF00196; GerE; 1.
CC PRINTS; PR00038; HTHLUXR.
CC ProDom; PD000307; HTH LuxR; 1.
CC SMART; SM00421; HTH_LUXR; 1.
CC PROSITE; PS00622; HTH_LUXR_FAMILY; FALSE NEG.
KW Transcription regulation; DNA-binding; Activator; Quorum sensing.
FT DNA BIND 193 212 H-T-H MOTIF (BY SIMILARITY).
SQ SEQUENCE 240 AA; 27253 MW; BF9E0C40564C6663 CRC64;

Query Match 2.2%; Score 9; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 0.66; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

QY 313 ILRLIQENQ 321
DB 4 ILRLIQENQ 12

RESULT 2
ID RSP4 STRPU STANDARD; PRT; 264 AA.
AC P46771;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 40S ribosomal protein SA (P40) (34/67 kDa laminin binding protein)
DE (Fragment).
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.

VANR VIBAN
ID VANR VIBAN STANDARD; PRT; 240 AA.
AC P74946;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transcriptional activator protein vanR.
GN VANR.
OS Vibrio anguillarum (Listonella anguillarum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Listonella.
OX NCBI_TaxID=55601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NB10 / Serotype O1;
RA MEDLINE=97284506; PubMed=9139920;
RX Milton D.L., Hardman A., Camara M., Chhabra S.R., Bycroft B.W.,
RA Stewart G.S.A.B., Williams P.;
RT "Quorum sensing in Vibrio anguillarum: characterization of the
RT vanI/vanR locus and identification of the autoinducer N-(3-
RT oxodecanoyl)-L-homoserine lactone.";
RL J. Bacteriol. 179:3004-3012(1997).
CC -|- SIMILARITY: PROBABLE TRANSCRIPTIONAL ACTIVATOR. BINDS TO AUTOINDUCER
CC MOLECULE ODHL.
CC -|- FUNCTION: BELONGS TO THE LUXR/UHPA FAMILY OF TRANSCRIPTIONAL
CC REGULATORS. AUTOINDUCER REGULATED SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U69677; AAC45213.1; -.
CC InterPro; IPR005143; Autoind bind.
CC InterPro; IPR000792; HTH LuxR.
CC Pfam; PF03472; Autoind_bind; 1.
CC Pfam; PF00196; GerE; 1.
CC PRINTS; PR00038; HTHLUXR.
CC ProDom; PD000307; HTH LuxR; 1.
CC SMART; SM00421; HTH_LUXR; 1.
CC PROSITE; PS00622; HTH_LUXR_FAMILY; FALSE NEG.
KW Transcription regulation; DNA-binding; Activator; Quorum sensing.
FT DNA BIND 193 212 H-T-H MOTIF (BY SIMILARITY).
SQ SEQUENCE 240 AA; 27253 MW; BF9E0C40564C6663 CRC64;

Query Match 2.2%; Score 9; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 0.66; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

QY 313 ILRLIQENQ 321
DB 4 ILRLIQENQ 12

RESULT 2
ID RSP4 STRPU STANDARD; PRT; 264 AA.
AC P46771;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 40S ribosomal protein SA (P40) (34/67 kDa laminin binding protein)
DE (Fragment).
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.

MEDLINE=96042034; PubMed=7589249;
Hung M., Rosenthal E.T., Boblett B., Benson S.;
"Characterization and localized expression of the laminin binding
protein/p40 (LBP/p40) gene during sea urchin development.";
Exp. Cell Res. 221:221-230(1995).
-|- SUBCELLULAR LOCATION: Cytoplasmic.
-|- SIMILARITY: BELONGS TO THE S2P FAMILY OF RIBOSOMAL PROTEINS.
-|- CAUTION: WAS ORIGINALLY THOUGHT TO BE A LAMININ RECEPTOR.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; U02369; AAA90976.1; -.
InterPro; IPR001865; Ribosomal_S2.
Pfam; PF00318; Ribosomal_S2; 1.
PROSITE; PS00962; RIBOSOMAL_S2_1; PARTIAL.
DR PROSITE; PS00963; RIBOSOMAL_S2_2; 1.
KW Ribosomal protein.
FT NON TER 1
SQ SEQUENCE 264 AA; 28141 MW; D0AB883C051CF859 CRC64;

Query Match 2.2%; Score 9; DB 1; Length 264;
Best Local Similarity 100.0%; Pred. No. 0.72; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

QY 99 PAAPVAPAA 107
DB 227 PAAPVAPAA 235

RESULT 3
PTGA CORGL
ID PTGA CORGL STANDARD; PRT; 674 AA.
AC Q45298;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE PTS system, glucose-specific IIAAC component (EIIABC-GLC) (Glucose-
DE permease IIAAC component) (Phosphotransferase enzyme II, ABC
DE component) (EC 2.7.1.69) (EII-GLC/EIII-GLC).
GN PTSG.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OX Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OX Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13869;
RA Yoon K.-H.;
DT "Cloning and nucleotide sequence of enzyme II of Brevibacterium
DT lactofermentum phosphotransferase system.";
DT Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
RL -|- FUNCTION: THIS IS A COMPONENT OF THE PHOSPHOENOLPYRUVATE-DEPENDENT
RL SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS). A MAJOR CARBOHYDRATE ACTIVE
RL -TRANSPORT SYSTEM. THE IID DOMAINS CONTAIN THE SUGAR BINDING SITE
RL AND THE TRANSMEMBRANE CHANNEL; THE IIA DOMAIN CONTAINS THE PRIMARY
RL PHOSPHORYLATION SITE (THE DONOR IS PHOSPHO-HPR); IIA TRANSFERS ITS
RL PHOSPHORYL GROUP TO THE IIB DOMAIN WHICH FINALLY TRANSFERS IT TO
RL THE SUGAR.
RL -|- CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar = protein
RL histidine + sugar phosphate.
RL -|- SUBCELLULAR LOCATION: Integral membrane protein.
RL -|- SIMILARITY: Contains 1 PTS EIIA domain.
RL -|- SIMILARITY: Contains 1 PTS EIIIB domain.
RL -|- SIMILARITY: Contains 1 PTS EIIC domain.
RL -|- SIMILARITY: Contains 1 PTS EIIC domain.
RL -----
RL This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; L18875; AAA22992.1; -;
 DR HSGP; P08837; IGIC.
 DR InterPro; IPR001127; PTS_EI1A.
 DR InterPro; IPR001196; PTS_EI1B.
 DR InterPro; IPR003352; PTS_EI1C.
 DR Pfam; PF00358; PTS_EI1A_1; 1.
 DR Pfam; PF00367; PTS_EI1B; 1.
 DR Pfam; PF02378; PTS_EI1C; 1.
 DR ProDom; PD002243; PTS_EI1A; 1.
 DR ProDom; PD001476; PTS_EI1B; 1.
 DR TIGRFAMs; TIGR00830; PTBA; 1.
 DR PROSITE; PS00371; PTS_EI1A_1; 1.
 DR PROSITE; PS01035; PTS_EI1B_CYS; 1.
 KW Phosphotransferase system; Sugar transport; Transferase;
 KW Phosphorylation; Transmembrane; Plasmid.
 FT DOMAIN 1 43
 FT DOMAIN ?
 FT DOMAIN 542 674
 FT TRANSMEM 126 146
 FT TRANSMEM 162 182
 FT TRANSMEM 193 213
 FT TRANSMEM 225 245
 FT TRANSMEM 260 280
 FT TRANSMEM 303 323
 FT TRANSMEM 344 364
 FT TRANSMEM 377 397
 FT TRANSMEM 409 429
 FT TRANSMEM 442 462
 FT TRANSMEM 493 513
 FT MOD RES 28 28
 FT MOD RES 594 594
 SQ SEQUENCE 674 AA; 71626 MW; 67A75AAF7E42FA2 CRC64;

Query Match 2.2%; Score 9; DB 1; Length 674;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 99 PAAPVAPAA 107
 |||||
 Db 490 PAAPVAPAA 498

RESULT 4

RL10_RHILO
 ID_RHILO STANDARD; PRT; 172 AA.
 AC Q98N68;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE 50S ribosomal protein L10.
 GN RPL7 OR MR0274
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RX MEDLINE=21082930; PubMed=11211968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti.";

RL DNA Res. 7:331-338(2000).
 CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AP002994; BAB47894.1; -;
 DR HAWAP; MF_00362; -; 1.
 DR InterPro; IPR001790; Ribosomal L10.
 DR InterPro; IPR002363; Ribosomal L10eub.
 DR Pfam; PF00466; Ribosomal L10; 1.
 DR PROSITE; PS01109; RIBOSOMAL_L10; 1.
 KW Ribosomal protein; Complete proteome.
 SQ SEQUENCE 172 AA; 17734 MW; 99AE3D33556AA5ED CRC64;

Query Match 2.0%; Score 8; DB 1; Length 172;
 Best Local Similarity 100.0%; Pred. No. 4.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 104 APAASVAR 111
 |||||
 Db 153 APAASVAR 160

RESULT 5

PDHR_ECOLI
 ID_PDHR_ECOLI STANDARD; PRT; 254 AA.
 AC P06957; O53381;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Pyruvate dehydrogenase complex repressor.
 GN PDHR OR GENA OR ACEC OR B0113 OR C0140 OR Z0123 OR ECS0117.
 OS Escherichia coli.
 OS Escherichia coli O6, and
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562, 217992, 83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=83209630; PubMed=6343085;
 RA Stephens P.E., Darlison M.G., Lewis H.M., Guest J.R.;
 RA "The pyruvate dehydrogenase complex of Escherichia coli K12.
 RT Nucleotide sequence encoding the pyruvate dehydrogenase component.";
 RL Eur. J. Biochem. 133:155-162(1983).
 RN [2]
 RP REVISIONS.
 RA Quail M.A.;
 RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=37426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glaesner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RA "The complete genome sequence of Escherichia coli K-12.";
 RT Science 277:1453-1474(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
 RX MEDLINE=22388234; PubMed=12471157;
 RA Weich R.A., Burland V., Plunkett G. III, Redford P., Roessch P.,
 RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,

RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic *Escherichia coli*";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
RN [5]

RC SEQUENCE FROM N.A.
RP STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,
Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7";
RL Nature 409:529-533 (2001).
RN [6]

RC SEQUENCE FROM N.A.
RP STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayaishi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RA "Complete genome sequence of enterohemorrhagic *Escherichia coli*
O157:H7 and genomic comparison with a laboratory strain K-12";
RL DNA Res. 8:11-22 (2001).
RN [7]

RC CHARACTERIZATION, AND VARIANT ACEC816.
RP STRAIN=K12;
RX MEDLINE=94085588; PubMed=8262214;
RA Haydon D.J., Quail M.A., Guest J.R.;
RT "A mutation causing constitutive synthesis of the pyruvate
dehydrogenase complex in *Escherichia coli* is located within the *pdhR*
gene";
RL FEBS Lett. 336:43-47 (1993).
RN [8]

CC -!- FUNCTION: TRANSCRIPTIONAL REPRESSOR FOR THE PYRUVATE DEHYDROGENASE
CC COMPLEX GENES ACEEF AND LPD
CC -!- SIMILARITY: BELONGS TO THE GNTR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; V01498; CAA24739.1; -;
DR EMBL; D26562; BAB96682.1; ALT_SEQ.
DR EMBL; A5000120; AAC73224.1; -;
DR EMBL; A5016755; AAN78636.1; ALT_INIT.
DR EMBL; A5005187; AAG54417.1; -;
DR EMBL; AF002550; BAB33540.1; -;
DR EMBL; S67363; AAB29356.1; -;
DR PIR; A64734; BVECA.
DR PIR; E85494; E85494.
DR PIR; E90643; E90643.
DR EcoGene; EGI1088; pdhR.
DR InterPro; IPR000524; HTH_Gntr.
DR Pfam; PF00392; gntr; 1.
DR PRINTS; PR00035; HTHGNTR.
DR SMART; SM00345; HTH GNTR; 1.
DR PROSITE; PS00043; HTH GNTR FAMILY; 1.
KW Transcription regulation; DNA-binding; Repressor; Complete proteome.
FT DNA_BIND 37 56 H-T-H MOTIF (BY SIMILARITY).
FT CONFLICT 118 118 R -> C (IN ACEC816; CONSTITUTIVELY
FT ACTIVATED).
SQ SEQUENCE 254 AA; 29425 MM; AF20245D303F751D CRC64;

Query Match 2.0%; Score 8; DB 1; Length 254;
Best Local Similarity 100.0%; Pred. No. 6.5;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 REAIQRLQ 366
Db 52 REAIQRLQ 59

RESULT 6
PDHR_SALTY STANDARD; PRT; 254 AA.
ID PDHR_SALTY STANDARD; PRT; 254 AA.
AC Q9L4H9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Pyruvate dehydrogenase complex repressor.
GN PDHR OR STM0351 OR STY0174 OR T0157.
OS *Salmonella typhimurium*, and
OC *Salmonella typhi*.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; *Salmonella*.
OX NCBI_TaxID=602, 601;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=SL1344;
RA Cano D., Casadesus J., Garcia-del Portillo F.;
RT "Characterization of a *Salmonella*-specific region located between *ampE*
and *aroP* genes";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
Waterston R., Wilson R.K.;
RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
LT2";
RL Nature 413:852-856 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
Churcher C., Mungall K.B., Bentley S.D., Holden M.T.G., Sebaihia M.,
Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant *Salmonella*
enterica serovar Typhi CT18";
RL Nature 413:848-852 (2001).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RA "Comparative genomics of *Salmonella enterica* serovar Typhi strains Ty2
and CT18";
RL J. Bacteriol. 185:2330-2337 (2003).
CC -!- FUNCTION: TRANSCRIPTIONAL REPRESSOR FOR THE PYRUVATE DEHYDROGENASE
CC COMPLEX GENES ACEEF AND LPD (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE GNTR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AJ242516; CAB98840.1; -;
 DR EMBL; AE008701; AAL19115.1; -;
 DR EMBL; AL627265; CAD01310.1; -;
 DR EMBL; AE016834; AAO67889.1; -;
 DR StyGene; SG77777; pdhr.
 DR InterPro; IPR000524; HTH_Gntr.
 DR Pfam; PF00392; gntcr; 1.
 DR PRINTS; PR00035; HTHGNTR.
 DR SMART; SMO0345; HTH_GNTR; 1.
 DR PROSITE; PS00043; HTH_GNTR_FAMILY; 1.
 KW Transcription regulation; DNA-binding; Repressor; Complete proteome.
 FT DNA_BIND 37 56 H-T-H MOTIF (BY SIMILARITY).
 FT CONFLICT 135 135 E -> K (IN REF. 1).
 SQ SEQUENCE 254 AA; 29411 MW; 990A66833052EAB3 CRC64;

Query Match 2.0%; Score 8; DB 1; Length 254;
 Best Local Similarity 100.0%; Pred. No. 6.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 REAIQRL 366
 |||||
 Db 52 REAIQRL 59

RESULT 7

ID STR_KLEPN STANDARD; PRT; 266 AA.
 AC P13082;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE Streptomycin 3'-kinase (EC 2.7.1.87) (Streptomycin 3'-
 DE phosphotransferase) (SPH).
 GN STR.
 OS Klebsiella pneumoniae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Klebsiella.
 OX NCBI_TaxID=573;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=85215465; PubMed=3889811;
 RA Mazodier P., Cossart P., Giraud E., Gasser F.;
 RT "Completion of the nucleotide sequence of the central region of Tn5
 RT confirms the presence of three resistance genes.";
 RL Nucleic Acids Res. 13:195-205(1985).
 CC -!- FUNCTION: THE AMINOGLYCOSIDE PHOSPHOTRANSFERASES ACHIEVE
 CC INACTIVATION OF THEIR ANTIBIOTIC SUBSTRATES BY PHOSPHORYLATION.
 CC -!- CATALYTIC ACTIVITY: ATP + streptomycin = ADP + streptomycin 3'-
 CC phosphate.
 CC -!- MISCELLANEOUS: THIS ENZYME IS ENCODED BY THE KANAMYCIN AND
 CC NEOMYCIN RESISTANCE TRANSPOSON TN5.
 CC -!- SIMILARITY: TO OTHER AMINOGLYCOSIDE PHOSPHOTRANSFERASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; U00004; AAA73392.1; -;
 DR EMBL; X01702; CAA25854.1; -;
 DR InterPro; IPR006748; APH_6_hur.
 DR Pfam; PF04655; APH_6_hur; 1.
 KW Antibiotic resistance; Transferase; Kinase; ATP-binding;
 KW Transposable element.
 FT ACT_SITE 154 154 BY SIMILARITY.
 SQ SEQUENCE 266 AA; 29066 MW; 30A9DA5D0E1A826C CRC64;

Query Match 2.0%; Score 8; DB 1; Length 266;
 Best Local Similarity 100.0%; Pred. No. 6.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 APAASVAR 111
 |||||
 Db 132 APAASVAR 139

RESULT 8

ID FRAH_ANASP STANDARD; PRT; 289 AA.
 AC P46017;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE FRAH protein.
 GN FRAH OR ALR1603.
 OS Anabaena sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 OX NCBI_TaxID=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bauer C.C., Scappino L., Haselkorn R.;
 RT "Bacterial subtracted cDNA libraries containing genes involved in the
 RT differentiation of vegetative cells to heterocysts allow a new twist
 RT on an old method of isolating developmental genes.";
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]

RP SEQUENCE FROM N.A.
 RX MEDLINE=21595285; PubMed=11759840;
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 RT cyanobacterium Anabaena sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).
 CC -!- FUNCTION: PUTATIVE HETEROCYST TO VEGETATIVE CELL CONNECTION.
 CC -!- SIMILARITY: Contains 1 FHA domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; U14553; AAA50356.1; -;
 DR EMBL; AP003586; BAB77969.1; -;
 DR PIR; AE2006; AE2006.
 DR InterPro; IPR000253; FHA.
 DR Pfam; PF00498; FHA; 1.
 DR SMART; SMO0240; FHA; 1.
 DR PROSITE; PS50006; FHA_DOMAIN; 1.
 KW Heterocyst; Zinc-finger; Complete proteome.
 FT ZN_FING 18 48 POTENTIAL.
 FT DOMAIN 67 172 PRO-RICH.
 FT DOMAIN 204 260 FHA.
 SQ SEQUENCE 289 AA; 29922 MW; 47C9091E50B39716 CRC64;

Query Match 2.0%; Score 8; DB 1; Length 289;
 Best Local Similarity 100.0%; Pred. No. 7.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 AAPAATVA 139
 |||||
 Db 56 AAPAATVA 63

RESULT 9

MGB3 HUMAN
 ID MGB3_HUMAN STANDARD; PRT; 346 AA.
 AC O15480; O75861;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Melanoma-associated antigen B3 (MAGE-B3 antigen).
 GN MAGEB3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98110575; PubMed=9441743;
 RA Lurquin C., de Smet C., Brasseur F., Muscatelli F., Martelange V.,
 RA de Plaen E., Brasseur R., Monaco A.P., Boon T.;
 RT "Two members of the human MAGEB gene family located in Xp21.3 are
 RT expressed in tumors of various histological origins";
 RL Genomics 46:397-408(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Muzny D., Arenson A.D., Adams C., Brundage E., Bunac C., Carvelli K.,
 RA Chacko J., Chen J., Di W., Ding Y., Dugan S., Durbin J., Forcum J.,
 RA Ganesh R., Garcia C., Goodman M., Gorrell J.H., Haywood M.,
 RA Hernandez J., Jackson L., Jin S., Kampal R., Karpathy S., Kovar C.,
 RA Leal B., Li Y., Lichtarge O., Liu W., Logan O., Lu J., Ly T.,
 RA Martinez C., Oswal G., Perez L., Raehid N.D., Rowland K., Savage L.,
 RA Scherer S.E., Shen H., Simon K., Stovall K., Timms K.M., Todd J.,
 RA Vo O., Williamson A., Worley K.C., Yu W., Chinault C., Nelson D.,
 RA Gibbs R.A.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- TISSUE SPECIFICITY: Expressed in testis.
 CC -1- SIMILARITY: Contains 1 MAGE domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U93163; AAC23618.1; -;
 CC DR EMBL; AC005185; AAD10636.1; -;
 CC DR Genew; HGNC:6810; MAGEB3.
 CC DR MIM; 300152; -;
 CC DR GO; GO:0008222; F.tumor antigen; TAS.
 CC DR InterPro; IPR002190; MAGE.
 CC DR Pfam; PF01454; MAGE; 1.
 CC DR PROSITE; PS00838; MAGE; 1.
 CC KW Antigen; Multigene family.
 CC FT DOMAIN 111 310 MAGE.
 CC FT CONFLICT 107 107 H -> R (IN REF. 2).
 CC FT CONFLICT 112 112 T -> I (IN REF. 2).
 CC SQ SEQUENCE 346 AA; 39179 MW; 4F9EABF6D813BA35 CRC64;
 Query Match 2.0%; Score 8; DB 1; Length 346;
 Best Local Similarity 100.0%; Pred. No. 8.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 75 SKAKASS 82
 Db 336 SKAKASS 343
 RESULT 10
 GUN1_USTMA STANDARD; PRT; 393 AA.
 ID GUN1_USTMA
 AC P54424;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Endoglucanase 1 precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase 1)
 DE (Cellulase 1) (EG 1).
 GN EGLI.
 OS Ustilago maydis (Smut fungus).
 OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
 OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
 OX NCBI_TaxID=5270;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96145728; PubMed=8590631;
 RA Schaevecker F., Wanner G., Kahmann R.;
 RT "Filament-specific expression of a cellulase gene in the dimorphic
 RT fungus Ustilago maydis";
 RL Biol. Chem. Hoppe-Seyler 376:617-625(1995).
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
 CC linkages in cellulose, lichenin and cereal beta-D-glucans.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: HYPHAL TIP.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN FILAMENTOUS DIKARYON.
 CC -1- PTM: MAY ALSO BE O-GLYCOSYLATED.
 CC -1- SIMILARITY: BELONGS TO CELLULASE FAMILY K (FAMILY 45 OF GLYCOSYL
 CC HYDROLASES).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; S81598; AAB36147.1; -;
 CC DR PIR; S59499; S59499.
 CC DR HSP; P43316; 2ENG.
 CC DR InterPro; IPR000334; Glyco_hydro 45.
 CC DR Pfam; PF02015; Glyco_hydro 45; 1.
 CC DR PROSITE; PS01140; GLYCOSYL_HYDROL_F45; 1.
 CC KW Cellulose degradation; Hydrolase; Glycosidase; Signal; Glycoprotein.
 CC FT SIGNAL 1 26 POTENTIAL.
 CC FT CHAIN 27 393 ENDOGLUCANASE 1.
 CC FT ACT_SITE 34 34 NUCLEOPHILE (BY SIMILARITY).
 CC FT ACT_SITE 152 152 PROTON DONOR (BY SIMILARITY).
 CC FT DOMAIN 270 385 ALA/GLY/SER-RICH.
 CC FT DOMBOHD 343 343 N-LINKED (GLCNAC...) (POTENTIAL).
 CC SQ SEQUENCE 393 AA; 39594 MW; 65C753C610CD6AD3 CRC64;
 Query Match 2.0%; Score 8; DB 1; Length 393;
 Best Local Similarity 100.0%; Pred. No. 9.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 259 PSGGSNPG 266
 Db 362 PSGGSNPG 369
 RESULT 11
 CAP_DICDI STANDARD; PRT; 464 AA.
 ID CAP_DICDI
 AC P54554;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Adenyl cyclase-associated protein (CAP).
 GN CAP.
 OS Dictyostellium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.
 OX NCBI_TaxID=4689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX3;
 RX MEDLINE=96228685; PubMed=8688557;

RA Gottwald U., Brokamp R., Karakesisoglou I., Schleicher M.,
RA Nogel A.A.;
RT Identification of a cyclase-associated protein (CAP) homologue in
RT Dictyostelium discoideum and characterization of its interaction with
RT actin.";
RL Mol. Biol. Cell 7:261-272(1996).
CC -1- FUNCTION: MAY HAVE A REGULATORY BIFUNCTIONAL ROLE. BINDS G-ACTIN
CC AND PIP2. INVOLVED IN MICROFILAMENT REORGANIZATION NEAR THE PLASMA
CC MEMBRANE IN A PIP2-REGULATED MANNER.
CC -1- SUBCELLULAR LOCATION: CELL MEMBRANE.
CC -1- DOMAIN: THE C-TERMINUS IS RESPONSIBLE FOR SEQUESTERING G-ACTIN.
CC THE N-TERMINUS IS REQUIRED FOR THE PIP2 MODULATION OF CAP
CC FUNCTION.
CC -1- SIMILARITY: BELONGS TO THE CAP FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U43027; AAB09713.1; -;
CC DictyDb; DD00012; -;
CC InterPro; IPR001837; CAP.
CC InterPro; IPR006599; CAP.
CC Pfam; PF01213; CAP; 1.
CC SMART; SMO0673; CAP; 2.
CC PROSITE; PS01088; CAP_1; 1.
CC PROSITE; PS01089; CAP_2; 1.
CC Membrane; Actin-binding.
CC KW DOMAIN 33 42 POLY-SER.
CC FT DOMAIN 217 253 PRO-RICH.
CC SQ SEQUENCE 464 AA; 49641 MW; 7ABDD05D8A0148B4 CRC64;

Query Match 2.0%; Score 8; DB 1; Length 464;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 PAAPVAPA 106
DB 234 PAAPVAPA 241

RESULT 12
BCB3 ACXY
ID BCB3 ACXY STANDARD; PRT; 804 AA.
AC Q9W822;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cyclic di-GMP binding protein precursor (Cellulose synthase regulatory
DE subunit) (Cellulose synthase protein B) (CDGPB).
GN BCSB1.
OS Acetobacter xylinus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Acetobacteraceae; Gluconacetobacter.
OX NCBI_TaxID=28448;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 7664 / IFO 13693;
RX MEDLINE=99310341; PubMed=10382968;
RA Umeda Y., Hirano A., Ishibashi M., Akiyama H., Onizuka T., Ikeuchi M.,
RA Inoue Y.;
RT Cloning of cellulose synthase genes from Acetobacter xylinum JCM
RT 7664: implication of a novel set of cellulose synthase genes.";
RL DNA Res. 6:109-115(1999).
CC -1- FUNCTION: Binds the cellulose synthase activator, bis-(3'-5')
CC cyclic diguanylic acid (c-di-GMP) (By similarity).
CC -1- PATHWAY: Bacterial cellulose biosynthesis.
CC -1- SUBUNIT: Tightly associated with the cellulose synthase catalytic
CC subunit (By similarity).

CC -1- SIMILARITY: BELONGS TO THE ACSB/BCSB FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB015802; BAA77586.1; -;
CC InterPro; IPR003920; CelSynth_B.
CC Pfam; PF031170; BCSB; 1.
CC PRINTS; PR01440; CELLSYNTHASEB.
CC Cellulose biosynthesis; Signal.
CC FT SIGNAL 1 18 POTENTIAL.
CC CHAIN 19 804 CYCLIC DI-GMP BINDING PROTEIN.
CC SQ SEQUENCE 804 AA; 85510 MW; 1888ADD82EFC50A3 CRC64;

Query Match 2.0%; Score 8; DB 1; Length 804;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 AAPVAPAA 107
DB 41 AAPVAPAA 48

RESULT 13
DPOL HSV11
ID DPOL HSV11 STANDARD; PRT; 1235 AA.
AC P04293;
DT 20-MAR-1987 (Rel. 04, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA polymerase (EC 2.7.7.7).
GN UL30.
OS Herpes simplex virus (type 1 / strain 17).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88274327; PubMed=2839594;
RA McGeoch D.J., Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,
RA McNab D., Perry L.J., Scott J.E., Taylor P.;
RT "The complete DNA sequence of the long unique region in the genome of
RT herpes simplex virus type 1.";
RL J. Gen. Virol. 69:1531-1574(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86067223; PubMed=2999714;
RA Quinn J.P., McGeoch D.J.;
RT "DNA sequence of the region in the genome of herpes simplex virus
RT type 1 containing the genes for DNA polymerase and the major DNA
RT binding protein.";
RL Nucleic Acids Res. 13:8143-8163(1985).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA] (N).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X14112; CAA32323.1; -;
CC EMBL; X03181; CAA26941.1; -;
CC PIR; A00715; DJBEVI.


```
DR PIR; C30085; DJBEH7.
DR InterPro; IPR0061172; DNA_pol_B.
DR InterPro; IPR0061134; DNA_pol_B_dom.
DR InterPro; IPR0061133; DNA_pol_B_exo.
DR Pfam; PF001136; DNA_pol_B; 1.
DR Pfam; PF03104; DNA_pol_B_exo; 1.
DR SMART; SM00486; POLBc; 1.
DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.
KW Transferase; DNA-directed DNA polymerase; DNA replication;
FT DNA-binding; Nuclear protein.
FT CONFLICT 114 126 VLRVSGGFWPRR -> APPRRVGRLLAA (IN
SQ SEQUENCE 1235 AA; 136420 MW; E8CD41D6EDED8343 CRC64;
REF. 2).

Query Match 2.0%; Score 8; DB 1; Length 1235;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 TLAQPAAP 102
DB 325 TLAQPAAP 332
|||||

RESULT 14
SGP2 CHRVI STANDARD; PRT; 124 AA.
AC OS2179;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Sulfur globule protein CV2 precursor.
GS SGPB OR SGP2.
OS Chromatium vinosum.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales;
OC Chromatiaceae; Allochromatium.
OX NCBI_TaxID=1049;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D / ATCC 17899 / DSM 180;
RX MEDLINE=98228197; PubMed=9560425;
RA Pattaragulanit K., Brune D.C., Trueper H.G., Dahl C.;
RT "Molecular genetic evidence for extracytoplasmic localization of
sulfur globules in Chromatium vinosum.";
RL Arch. Microbiol. 169:434-444(1998).
RN [2]
RP SEQUENCE OF 23-72, AND CHARACTERIZATION.
RC STRAIN=D / ATCC 17899 / DSM 180;
RX MEDLINE=96009136; PubMed=7575095;
RA Brune D.C.;
RT "Isolation and characterization of sulfur globule proteins from
Chromatium vinosum and Thiocapsa roseopersicina.";
RL Arch. Microbiol. 163:191-199(1995)
CC -1- FUNCTION: STRUCTURAL PROTEIN OF THE SULFUR GLOBULES, WHICH ARE
INTRACELLULAR GLOBULES THAT SERVE FOR SULFUR STORAGE IN PURPLE
SULFUR BACTERIA.
CC -1- SUBUNIT: THE PROTEIN ENVELOPE OF THE SULFUR GLOBULES ARE COMPOSED
OF THREE DIFFERENT PROTEINS; CVL, CV2 AND CV3.
CC -1- MASS SPECTROMETRY: MW=10451; METHOD=MALDI; RANGE=22-124.
CC -1- SIMILARITY: TO C.VINOSUM CV1 AND T.ROSEOPERSICINA TR0.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF025952; AAB91547.1; -.
KW Structural protein; Signal.
FT SIGNAL 1 22
FT CHAIN 23 124 Sulfur globule protein CV2.
SQ SEQUENCE 124 AA; 12607 MW; 2220B28FFDC2525F CRC64;
```

```
Query Match 1.7%; Score 7; DB 1; Length 124;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 APVAPAA 107
DB 111 APVAPAA 117
|||||

RESULT 15
CU24 ARADI STANDARD; PRT; 126 AA.
AC P80516;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adult-specific rigid cuticular protein 12.4 (ACP 12.4).
OS Araneus diadematus (Spider).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Araneomorphae; Entelegynae; Araneidae; Araneidae; Araneus.
OX NCBI_TaxID=45920;
RN [1]
RP SEQUENCE.
RC TISSUE=Cuticle;
RX MEDLINE=97166616; PubMed=9014336;
RA Norup T., Berg T., Stenholm H., Andersen S.O., Hoejrup P.;
RT "Purification and characterization of five cuticular proteins from
the spider Araneus diadematus";
RL Insect Biochem. Mol. Biol. 26:907-915(1996).
CC -1- FUNCTION: COMPONENT OF THE RIGID CUTICLE OF THE SPIDER.
CC -1- MASS SPECTROMETRY: MW=12420.5; METHOD=Electrospray.
CC -1- SIMILARITY: Contains 1 cuticle consensus domain.
DR InterPro; IPR000618; Insect cuticle.
DR Pfam; PF00379; Chitin bind_4; 1.
DR PROSITE; PS00233; CUTICLE; 1.
KW Structural protein; Cuticle.
FT DOMAIN 73 83 POLY-ALA.
SQ SEQUENCE 126 AA; 12421 MW; F9633E26A877F6DE CRC64;

Query Match 1.7%; Score 7; DB 1; Length 126;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 PAAPVAP 105
DB 92 PAAPVAP 98
|||||

RESULT 16
RISB LACLA STANDARD; PRT; 154 AA.
AC Q9CGU6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 6,7-dimethyl-8-ribityllumazine synthase [EC 2.5.1.9] (DMRL synthase)
[umazine synthase] (Riboflavin synthase beta chain).
GN RISB OR LL0996.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Manger S., Jaillon O., Malarre K.,
Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
lactis ssp. lactis IL1403";
RL Genome Res. 11:731-753(2001).
CC -1- FUNCTION: Riboflavin synthase is a bifunctional enzyme complex
catalyzing the formation of riboflavin from 5-amino-6-(1'-D)-
```

```

CC ribityl-amino-2,4-(1H,3H)-pyrimidinone and L-3,4-dihydroxy-2-
CC butanone-4-phosphate via 6,7-dimethyl-8-lumazine. The beta subunit
CC catalyzes the condensation of 5-amino-6-(1'-D)-ribityl-amino-
CC 2,4(1H,3H)-pyrimidinone with L-3,4-dihydroxy-2-butanone-4-
CC phosphate yielding 6,7-dimethyl-8-lumazine (By similarity).
CC -!- CATALYTIC ACTIVITY: 2,6,7-dimethyl-8-(1-D-ribityl)lumazine =
CC riboflavin + 4-(1-D-ribitylamino)-5-amino-2,6-dihydroxypyrimidine.
CC -!- PATHWAY: Riboflavin biosynthesis; last step.
CC -!- SIMILARITY: Belongs to the DMRL synthase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE006333; AAK05094.1; -.
DR PIR; D86749; D86749.
DR HSPP; P11998; IRVV.
DR HAWAP; MF_00178; -.
DR InterPro; IPR002180; DMRL synthase.
DR Pfam; PF00885; DMRL synthase; 1.
DR ProDom; PD003664; DMRL synthase; 1.
DR TIGRFAMs; TIGR00114; ribH; 1.
DR Riboflavin biosynthesis; Transferase; Complete proteome.
KW RIBOFLAVIN biosynthesis; Transferase; Complete proteome.
SQ SEQUENCE 154 AA; 16696 MW; 2E06CD2C2979A3BF CRC64;
-----
Query Match 1.7%; Score 7; DB 1; Length 154;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 273 SGALDAL 279
Db 32 SGALDAL 38
|||||
RESULT 17
RIB2 PHOLE STANDARD; PRT; 156 AA.
AC Q93E92;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 6,7-dimethyl-8-ribityllumazine synthase (EC 2.5.1.9) (DMRL synthase)
DE (Lumazine synthase) (Riboflavin synthase beta chain).
GN RIBH.
OS Photobacterium leiognathi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Photobacterium.
OX NCBI_TaxID=658;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PL741;
RA MEDLINE=21290647; PubMed=11396941;
RA Lin J.-W., Chao Y.-F., Weng S.-F.;
RA "Riboflavin synthesis genes ribE, ribB, ribH, ribA reside in the lux
RT operon of Photobacterium leiognathi";
RL Biochem. Biophys. Res. Commun. 284:587-595(2001).
CC -!- FUNCTION: Riboflavin synthase is a bifunctional enzyme complex
CC catalyzing the formation of riboflavin from 5-amino-6-(1'-D)-
CC ribityl-amino-2,4(1H,3H)-pyrimidinone and L-3,4-dihydroxy-2-
CC butanone-4-phosphate via 6,7-dimethyl-8-lumazine. The beta subunit
CC catalyzes the condensation of 5-amino-6-(1'-D)-ribityl-amino-
CC 2,4(1H,3H)-pyrimidinone with L-3,4-dihydroxy-2-butanone-4-
CC phosphate yielding 6,7-dimethyl-8-lumazine (By similarity).
CC -!- CATALYTIC ACTIVITY: 2,6,7-dimethyl-8-(1-D-ribityl)lumazine =
CC riboflavin + 4-(1-D-ribitylamino)-5-amino-2,6-dihydroxypyrimidine.
CC -!- PATHWAY: Riboflavin biosynthesis; last step.
CC -!- SIMILARITY: Belongs to the DMRL synthase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF364106; AAK83294.1; -.
DR HAWAP; MF_00178; -.
DR InterPro; IPR002180; DMRL synthase.
DR Pfam; PF00885; DMRL synthase; 1.
DR ProDom; PD003664; DMRL synthase; 1.
DR TIGRFAMs; TIGR00114; ribH; 1.
DR Riboflavin biosynthesis; Transferase.
KW RIBOFLAVIN biosynthesis; Transferase.
SQ SEQUENCE 156 AA; 16423 MW; B53E2727C6BCD0A5 CRC64;
-----
Query Match 1.7%; Score 7; DB 1; Length 156;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 273 SGALDAL 279
Db 32 SGALDAL 38
|||||
RESULT 18
RRF TREPA STANDARD; PRT; 183 AA.
AC O83613;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ribosome recycling factor (Ribosome releasing factor) (RRF).
DE FRR OR TP0604.
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RA MEDLINE=9832770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uettersack T.,
RA McDonald L., Artiaach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RA "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";
RL Science 281:375-388(1998).
CC -!- FUNCTION: RESPONSIBLE FOR THE RELEASE OF RIBOSOMES FROM MESSENGER
CC RNA AT THE TERMINATION OF PROTEIN BIOSYNTHESIS. MAY INCREASE THE
CC EFFICIENCY OF TRANSLATION BY RECYCLING RIBOSOMES FROM ONE ROUND OF
CC TRANSLATION TO ANOTHER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE RRF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001235; AAC65577.1; -.
DR PIR; F71304; F71304.
DR HSPP; Q9X1B9; 1DD5.
DR TIGR; TP0604; -.
DR HAWAP; MF_00040; -.
DR InterPro; IPR002661; RRF.
DR Pfam; PF01765; RRF; 1.

```



```

QY      26 VKRIIET 32
DB      179 VKRIIET 185

RESULT 21
OLP1_LYCES
ID_OLP1_LYCES STANDARD; PRT; 252 AA.
AC Q41350;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Osmotin-like protein precursor.
OS Osmotin-like protein precursor.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamiids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. VF36;
RX MEDLINE=97128324; PubMed=8972917;
RA Chen R., Wang F., Smith A.G.;
RT "A flower-specific gene encoding an osmotin-like protein from
RT Lycopersicon esculentum.";
RL Gene 179:301-302(1996).
RN [2]
RP SEQUENCE OF 25-33.
RA Robertson D., Mitchell G.P., Gilroy J.S., Gerrish C., Bolwell G.P.,
RA Slabas A.R.;
RL Submitted (JAN-1997) to the SWISS-PROT data bank.
CC -1- SUBCELLULAR LOCATION: Cell wall (potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL; L76632; AAB41124.1; -
DR PIR; JC5237; JC5237.
DR HSPG; P25871; LAUN.
DR InterPro; IPR001938; Thaumatin.
DR Pfam; PF00314; thaumatin; 1.
DR PRINTS; PR00347; THAUMATIN.
DR ProDom; PD003321; Thaumatin; 1.
DR SMART; SM00205; THN; 1.
DR PROSITE; PS00316; THAUMATIN; 1.
KW Cell wall; Signal.
FT SIGNAL 1 24
FT CHAIN 25 252 OSMOTIN-LIKE PROTEIN.
SQ SEQUENCE 252 AA; 27265 MW; 36031901A4F6E190 CRC64;

Query Match 1.7%; Score 7; DB 1; Length 252;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      92 APATIAQ 98
DB      113 APATIAQ 119

RESULT 22
HIS4_SYNY3
ID HIS4_SYNY3 STANDARD; PRT; 256 AA.
AC P74561;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino]

```

```

DE imidazole-4-carboxamide isomerase (EC 5.3.1.16)
DE (Phosphoribosylformimino-5-aminoimidazole carboxamide ribotide
DE isomerase).
GN HISA OR SLR0652.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kareko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirosewa M., Sugiyura M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
RA Okumura S., Shimpou S., Takeuchi C., Wada T., Watanabe A.,
RA Yamada M., Yasuda M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -1- CATALYTIC ACTIVITY: 1-(5-phosphoribosyl)-5-[(5-
CC phosphoribosylamino)methylideneamino]imidazole-4-carboxamide = 5-
CC [(5-phospho-1-deoxyribulos-1-ylamino)methylideneamino]-1-(5-
CC phosphoribosyl)imidazole-4-carboxamide.
CC -1- PATHWAY: Histidine biosynthesis; fourth step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: BELONGS TO THE HISA / HISP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL; D90916; BAA18668.1; -
DR PIR; S76756; S76756.
DR HAMAP; MF 01014; -; 1.
DR InterPro; IPR003009; FWN enzyme.
DR InterPro; IPR006063; His4.
DR InterPro; IPR006062; His_biosynth.
DR Pfam; PF00977; His_biosynth; 1.
DR TIGRFAMs; TIGR00007; TIGR00007; 1.
DR Isomerase; Histidine biosynthesis; Complete proteome.
KW Isomerase; Histidine biosynthesis; Complete proteome.
SQ SEQUENCE 256 AA; 27350 MW; 2759585808P7F439 CRC64;

Query Match 1.7%; Score 7; DB 1; Length 256;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      276 LDALROL 282
DB      179 LDALROL 185

RESULT 23
TNR9_MOUSE
ID_TNR9_MOUSE STANDARD; PRT; 256 AA.
AC P20334;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tumor necrosis factor receptor superfamily member 9 precursor (4-1BB
DE ligand receptor) (T-cell antigen 4-1BB) (CD137 antigen).
GN TNFRSF9 OR ILA OR LY63 OR CD137 OR CD157.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89184547; PubMed=2784565;
RA Kwon B.S., Weissman S.M.;

```

```
RT "cDNA sequences of two inducible T-cell genes.";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:1963-1967(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/c;  
RX MEDLINE=94179805; PubMed=8133039;  
RA Kwon B.S., Kozak C.A., Kim K.K., Pickard R.T.;  
RT "Genomic organization and chromosomal localization of the T-cell  
antigen 4-1BB.";  
RL J. Immunol. 152:2256-2262(1994).  
RN [3]  
RP CHARACTERIZATION, AND SEQUENCE OF 25-29.  
RX MEDLINE=93139510; PubMed=7678621;  
RA Pollok K.E., Kim Y.-J., Zhou Z., Hurtado J., Kin K.K., Pickard R.T.,  
RT Kwon B.S.;  
RT "Inducible T cell antigen 4-1BB. Analysis of expression and  
function.";  
RL J. Immunol. 150:771-781(1993).  
CC -|- FUNCTION: Receptor for TNFSP14/4-1BBL. Possibly active during T  
cell activation.  
CC -|- SUBUNIT: PRINCIPALLY AN HOMODIMER, BUT ALSO FOUND AS A MONOMER.  
CC -|- ASSOCIATES WITH F56-LCK. Interacts with TRAF1, TRAF2 AND TRAF3 (By  
similarity).  
CC -|- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -|- TISSUE SPECIFICITY: EXPRESSED ON THE SURFACE OF ACTIVATED T-CELLS.  
CC -|- INDUCTION: OPTIMAL BY PMA AND IONOMYCIN.  
CC -|- SIMILARITY: Contains 4 TNFR-Cys repeats.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL; J04492; AAA40167.1; -;  
DR EMBL; U02567; AAA93113.1; -;  
DR PIR; B32393; B32393.  
DR PDB; 1D0J; 26-SEP-01.  
DR MGD; MGI:1101059; Tnfrsf9.  
DR InterPro; IPR001368; TNFR_C6.  
DR Pfam; PF00020; TNFR_C6; 1.  
DR SMART; SM00208; TNFR; 2.  
DR PROSITE; PS00652; TNFR_NGFR_1; 1.  
DR PROSITE; PS00050; TNFR_NGFR_2; FALSE NEG.  
KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal; 3D-structure.  
FT SIGNAL 1 24  
FT CHAIN 25 256 TUMOR NECROSIS FACTOR RECEPTOR  
FT FT SUPERFAMILY MEMBER 9.  
FT DOMAIN 25 187 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 188 208 POTENTIAL.  
FT DOMAIN 209 256 CYTOPLASMIC (POTENTIAL).  
FT REPEAT 17 45 TNFR-CYS 1.  
FT REPEAT 46 85 TNFR-CYS 2.  
FT REPEAT 86 117 TNFR-CYS 3.  
FT REPEAT 118 159 TNFR-CYS 4.  
FT DISULFID 28 37 BY SIMILARITY.  
FT DISULFID 31 44 BY SIMILARITY.  
FT DISULFID 47 61 BY SIMILARITY.  
FT DISULFID 64 77 BY SIMILARITY.  
FT DISULFID 67 85 BY SIMILARITY.  
FT DISULFID 87 93 BY SIMILARITY.  
FT DISULFID 98 105 BY SIMILARITY.  
FT DISULFID 101 116 BY SIMILARITY.  
FT DISULFID 119 133 BY SIMILARITY.  
FT DISULFID 139 158 BY SIMILARITY.  
FT CARBOHYD 128 128 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 256 AA; 27598 MW; 93A10D03C60813C4 CRC64;  
  
Query Match 1.7%; Score 7; DB 1; Length 256;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
RT "CDNA sequences of two inducible T-cell genes.";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:1963-1967(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/c;  
RX MEDLINE=94179805; PubMed=8133039;  
RA Kwon B.S., Kozak C.A., Kim K.K., Pickard R.T.;  
RT "Genomic organization and chromosomal localization of the T-cell  
antigen 4-1BB.";  
RL J. Immunol. 152:2256-2262(1994).  
RN [3]  
RP CHARACTERIZATION, AND SEQUENCE OF 25-29.  
RX MEDLINE=93139510; PubMed=7678621;  
RA Pollok K.E., Kim Y.-J., Zhou Z., Hurtado J., Kin K.K., Pickard R.T.,  
RT Kwon B.S.;  
RT "Inducible T cell antigen 4-1BB. Analysis of expression and  
function.";  
RL J. Immunol. 150:771-781(1993).  
CC -|- FUNCTION: Receptor for TNFSP14/4-1BBL. Possibly active during T  
cell activation.  
CC -|- SUBUNIT: PRINCIPALLY AN HOMODIMER, BUT ALSO FOUND AS A MONOMER.  
CC -|- ASSOCIATES WITH F56-LCK. Interacts with TRAF1, TRAF2 AND TRAF3 (By  
similarity).  
CC -|- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -|- TISSUE SPECIFICITY: EXPRESSED ON THE SURFACE OF ACTIVATED T-CELLS.  
CC -|- INDUCTION: OPTIMAL BY PMA AND IONOMYCIN.  
CC -|- SIMILARITY: Contains 4 TNFR-Cys repeats.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL; J04492; AAA40167.1; -;  
DR EMBL; U02567; AAA93113.1; -;  
DR PIR; B32393; B32393.  
DR PDB; 1D0J; 26-SEP-01.  
DR MGD; MGI:1101059; Tnfrsf9.  
DR InterPro; IPR001368; TNFR_C6.  
DR Pfam; PF00020; TNFR_C6; 1.  
DR SMART; SM00208; TNFR; 2.  
DR PROSITE; PS00652; TNFR_NGFR_1; 1.  
DR PROSITE; PS00050; TNFR_NGFR_2; FALSE NEG.  
KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal; 3D-structure.  
FT SIGNAL 1 24  
FT CHAIN 25 256 TUMOR NECROSIS FACTOR RECEPTOR  
FT FT SUPERFAMILY MEMBER 9.  
FT DOMAIN 25 187 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 188 208 POTENTIAL.  
FT DOMAIN 209 256 CYTOPLASMIC (POTENTIAL).  
FT REPEAT 17 45 TNFR-CYS 1.  
FT REPEAT 46 85 TNFR-CYS 2.  
FT REPEAT 86 117 TNFR-CYS 3.  
FT REPEAT 118 159 TNFR-CYS 4.  
FT DISULFID 28 37 BY SIMILARITY.  
FT DISULFID 31 44 BY SIMILARITY.  
FT DISULFID 47 61 BY SIMILARITY.  
FT DISULFID 64 77 BY SIMILARITY.  
FT DISULFID 67 85 BY SIMILARITY.  
FT DISULFID 87 93 BY SIMILARITY.  
FT DISULFID 98 105 BY SIMILARITY.  
FT DISULFID 101 116 BY SIMILARITY.  
FT DISULFID 119 133 BY SIMILARITY.  
FT DISULFID 139 158 BY SIMILARITY.  
FT CARBOHYD 128 128 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 256 AA; 27598 MW; 93A10D03C60813C4 CRC64;  
  
Query Match 1.7%; Score 7; DB 1; Length 256;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 332 PEGGPGG 338  
DB 175 PEGGPGG 181  
  
RESULT 24  
LIGE PSEPA  
ID LIGE PSEPA STANDARD; PRT; 280 AA.  
AC P27457;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE Beta-etherase (Beta-aryl ether cleaving enzyme).  
GN LIGE  
OS Pseudomonas paucimobilis (Sphingomonas paucimobilis).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;  
OC Sphingomonadaceae; Sphingomonas.  
OX NCBI_TaxID=13689;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-10.  
RC STRAIN=SYK-6;  
RX MEDLINE=92078104; PubMed=1744051;  
RA Masai E., Katayama Y., Kawai S., Nishikawa S., Yamasaki M.,  
RA Morohoshi N.;  
RT "Cloning and sequencing of the gene for a Pseudomonas paucimobilis  
enzyme that cleaves beta-aryl ether.";  
RL J. Bacteriol. 173:7950-7955(1991).  
CC -|- FUNCTION: ABLE TO DEGRADE VARIOUS DIMERIC LIGNIN COMPOUNDS.  
CC CATALYZES THE UNIQUE AND REDUCTIVE CLEAVAGE OF ARYLGLYCEROL-  
BETA-ARYL ETHER.  
CC -|- SUBCELLULAR LOCATION: Inner membrane-associated.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL; M69107; AAA25878.1; -;  
DR EMBL; D11473; BAA02032.1; -;  
DR PIR; A43749; A43749.  
DR InterPro; IPR004046; GST_Cterm.  
DR InterPro; IPR004045; GST_Nterm.  
DR Pfam; PF00043; GST_C; 1.  
DR Pfam; PF02798; GST_N; 1.  
KW Hydrolase; Lignin degradation; Inner membrane.  
FT INIT MET 0  
FT SEQUENCE 280 AA; 31939 MW; 7A5705D2A29D9C5B CRC64;  
  
Query Match 1.7%; Score 7; DB 1; Length 280;  
Best Local Similarity 100.0%; Pred. No. 66;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 107 ASVARTP 113  
DB 198 ASVARTP 204  
  
RESULT 25  
NAR4 HUMAN  
ID NAR4 HUMAN STANDARD; PRT; 314 AA.  
AC Q93070; Q9BZ50; Q9BZ51; Q9HB06;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Ecto-ADP-ribosyltransferase 4 precursor (EC 2.4.2.31) (NAD(P)(+)-  
arginine ADP-ribosyltransferase 4) (Mono (ADP-ribosyl)transferase 4)  
DE (Dombrock blood group carrier molecule).  
GN DO OR ART4 OR DOK1.
```

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANT DO(B).
 RX MEDLINE=20458812; PubMed=11001920;
 RA Gubin A.N., Njoroge J.M., Wojda U., Pack S.D., Rios M., Reid M.E.,
 RA Miller J.L.;
 RT "Identification of the Dombrock blood group glycoprotein as a
 RT polymorphic member of the ADP-ribosyltransferase gene family.";
 RL Blood 96:2621-2627(2000).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANTS DO(B); JO(A); HY1 AND HY2.
 RX MEDLINE=21893830; PubMed=11896313;
 RA Rios M., Hue-Roye K., Oyen R., Miller J., Reid M.E.;
 RT "Insights into the Holley- and Joseph- phenotypes.";
 RL Transfusion 42:52-58(2002).
 RN [3]
 RP SEQUENCE OF 48-314 FROM N.A.
 RX MEDLINE=97224466; PubMed=9119374;
 RA Koch-Noite F., Haag F., Braren R., Kuehl M., Hoovers J.,
 RA Balasubramanian S., Bazan F., Thiele H.-G.;
 RT "Two novel human members of an emerging mammalian gene family related
 RT to mono-ADP-ribosylating bacterial toxins.";
 RL Genomics 39:370-376(1997).
 RN [4]
 RP SEQUENCE OF 49-284 FROM N.A., AND VARIANT DO(B).
 RX MEDLINE=21412425; PubMed=11520417;
 RA Wu G.-G., Jin S.-Z., Deng Z.-H., Zhao T.-M.;
 RT "Polymerase chain reaction with sequence-specific primers-based
 RT genotyping of the human Dombrock blood group D01 and D02 alleles and
 RT the DO gene frequencies in Chinese blood donors.";
 RL Vox Sang. 81:49-51(2001).
 CC -!- CATALYTIC ACTIVITY: NAD(+) + L-arginine = nicotinamide + N(2)-
 CC (ADP-D-ribosyl)-L-arginine.
 CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN SPLEEN.
 CC -!- POLYMORPHISM: DO is responsible for the Dombrock blood group
 CC system. The molecular basis of the Do(a)/Do(b) blood group antigen
 CC is a single variation in position 265; Asn-265 corresponds to
 CC Do(a) and Asp-265 to Do(b). It is also responsible for the
 CC antigens Gregory [Gy(a)], Holley [Hy] and Joseph [Jo(a)].
 CC -!- SIMILARITY: BELONGS TO THE ARG-SPECIFIC ADP-RIBOSYLTRANSFERASE
 CC FAMILY.
 CC -!- DATABASE: NAME=Blood group antigen mutation database;
 CC NOTE=Dombrock (DO) blood group system;
 CC WWW="http://www.bioc.aecom.yu.edu/bgmut/dombrock.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF230204; AAG17845.1; -;
 DR EMBL; AF382213; AAM21462.1; -;
 DR EMBL; AF382211; AAM21462.1; JOINED.
 DR EMBL; AF382212; AAM21462.1; JOINED.
 DR EMBL; AF382216; AAM21464.1; -;
 DR EMBL; AF382214; AAM21464.1; JOINED.
 DR EMBL; AF382215; AAM21464.1; JOINED.
 DR EMBL; AF382219; AAM21465.1; -;
 DR EMBL; AF382217; AAM21465.1; JOINED.
 DR EMBL; AF382218; AAM21465.1; JOINED.
 DR EMBL; AF382222; AAM21466.1; -;
 DR EMBL; AF382220; AAM21466.1; JOINED.
 DR EMBL; AF382221; AAM21466.1; JOINED.
 DR EMBL; AF382225; AAM21467.1; -;
 DR EMBL; AF382223; AAM21467.1; JOINED.
 DR EMBL; AF382224; AAM21467.1; JOINED.

DR EMBL; X95926; CAA65095.1; -;
 DR EMBL; AF340233; AAK11274.1; -;
 DR EMBL; AF340234; AAK11275.1; -;
 DR Genew; HGNC:726; DO.
 DR MIM; 603087; -;
 DR MIM; 110600; -;
 DR GO; GO:0016020; C:membrane; NAS.
 DR GO; GO:0003956; F:NAD(P)+-arginine ADP-ribosyltransferase act. .; NAS.
 DR GO; GO:0006525; P:arginine metabolism; NAS.
 DR GO; GO:0006471; P:protein amino acid ADP-ribosylation; NAS.
 DR InterPro; IPR000768; ART_family.
 DR Pfam; PF01129; ART; 1.
 DR PRINTS; PR00970; RIBTRNSFRASE.
 DR PROSITE; PS01291; ART; 1.
 KW Transferase; Glycosyltransferase; Glycoprotein; NAD; GPI-anchor;
 KW Signal; Polymorphism; Blood group antigen.
 FT SIGNAL 1 46
 FT CHAIN 47 285
 FT PROPEP 286 314
 FT ACT SITE 251 251
 FT LIPID 285 285
 FT CARBOHYD 114 114
 FT CARBOHYD 178 178
 FT CARBOHYD 222 222
 FT CARBOHYD 257 257
 FT CARBOHYD 274 274
 FT VARIANT 108 108
 FT VARIANT 117 117
 FT VARIANT 265 265
 FT VARIANT 300 300
 FT CONFLICT 48 48
 FT CONFLICT 116 116
 FT SEQUENCE 314 AA; 35877 MW; 174913890C9D158F CRC64;
 SQ
 Query Match 1.7%; Score 7; DB 1; Length 314;
 Best Local Similarity 100.0%; Pred. No. 73;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 107 ASVARTP 113
 Db 141 ASVARTP 147
 RESULT 26
 ID NAR4_PANTR STANDARD; PRT; 314 AA.
 AC Q9SNE0;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ecto-ADP-ribosyltransferase 4 precursor (EC 2.4.2.31) (NAD(P) (+)-
 DE arginine ADP-ribosyltransferase 4) (Mono(ADP-ribosyl)transferase 4)
 DE (Dombrock molecule 1).
 GN DO OR ART4.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Rios M., Reid M.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: NAD(+) + L-arginine = nicotinamide + N(2)-
 CC (ADP-D-ribosyl)-L-arginine.
 CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By
 CC similarity).
 CC -!- SIMILARITY: BELONGS TO THE ARG-SPECIFIC ADP-RIBOSYLTRANSFERASE
 CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF373017; AAK53821.1; --
 DR EMBL; AF373016; AAK53821.1; JOINED.
 DR EMBL; AF374727; AAK53821.1; JOINED.
 DR GO; GO:0016020; C:membrane; ISS.
 DR GO; GO:0003956; F:NAD(P)+-arginine ADP-ribosyltransferase act. . .; ISS.
 DR GO; GO:0006525; P:arginine metabolism; ISS.
 DR GO; GO:0006471; P:protein amino acid ADP-ribosylation; ISS.
 DR InterPro; IPR000768; ART_family.
 DR Pfam; PF01129; ART; 1.
 DR PRINTS; PR00970; RIBTRNSFRASE.
 DR PROSITE; PS01291; ART; 1.
 DR Transferase; Glycosyltransferase; Glycoprotein; NAD; GPI-anchor;
 KW Signal.
 FT SIGNAL 1 46 POTENTIAL.
 FT CHAIN 47 285 ECTO-ADP-RIBOSYLTRANSFERASE 4.
 FT PROPEP 286 314 REMOVED IN MATURE FORM (POTENTIAL).
 FT ACT SITE 251 251 BY SIMILARITY.
 FT LIPID 285 285 GPI-ANCHOR (POTENTIAL).
 FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 222 222 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 257 257 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 314 AA; 35861 MW; 1CD76A650B87A423 CRC64;

Query Match 1.7%; Score 7; DB 1; Length 314;
 Best Local Similarity 100.0%; Pred.No. 73;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 ASVARTP 113
 DB 141 ASVARTP 147

RESULT 27
 CAHC_HORVU STANDARD; PRT; 324 AA.
 AC P40880;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DE Carbonic anhydrase, chloroplast precursor (EC 4.2.1.1) (Carbonate
 DE dehydratase).
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Hordeum.
 OX NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95303975; PubMed=7784519;
 RA Bracey M.H., Bartlett S.G.;
 RT "Sequence of a cDNA encoding carbonic anhydrase from barley.";
 RL Plant Physiol. 108:433-434(1995).
 CC -1- FUNCTION: Reversible hydration of carbon dioxide.
 CC -1- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.
 CC -1- SUBUNIT: Homohexamer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Chloroplast stroma (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE PLANT AND PROKARYOTIC CARBONIC
 CC ANHYDRASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; U36959; AAC41656.1; --
 DR PIR; T04478; T04478.
 DR InterPro; IPR001765; Prok_COanhd.
 DR Pfam; PF00484; Pro CA; 1.
 DR PROSITE; PS00704; PROK_CO2 ANHYDRASE 1; 1.
 DR PROSITE; PS00705; PROK_CO2 ANHYDRASE 2; 1.
 KW Lyase; Zinc; Chloroplast; Transit peptide.
 FT TRANSIT 1 ? CHLOROPLAST (POTENTIAL).
 FT CHAIN ? 324 CARBONIC ANHYDRASE.
 SQ SEQUENCE 324 AA; 35074 MW; 026F541956F06E75 CRC64;

Query Match 1.7%; Score 7; DB 1; Length 324;
 Best Local Similarity 100.0%; Pred.No. 76;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 AAPVAPA 106
 DB 115 AAPVAPA 121

RESULT 28
 G3P_HALN1 STANDARD; PRT; 335 AA.
 ID G3P_HALN1
 AC Q9HSS7;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.59) (GAPDH)
 DE (NAD(P)-dependent glyceraldehyde-3-phosphate dehydrogenase).
 DE GAP OR GAPB OR VNG0095G.
 GN Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
 OS Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 OC Halobacteriaceae; Halobacterium.
 OX NCBI_TaxID=64091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20504483; PubMed=11016950;
 RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
 RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
 RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
 RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudis J.L., Jung K.-H.,
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Ebbardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
 RT "Genome sequence of Halobacterium species NRC-1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
 CC NAD(P)(+) = 3-phospho-D-glyceroyl phosphate + NAD(P)H.
 CC -1- PATHWAY: Second phase of glycolysis; first step.
 CC -1- SUBUNIT: Homotetramer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
 CC dehydrogenase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF004977; AAG18725.1; --
 DR PIR; A84170; A84170.
 DR HSSP; P10618; 1CF2.
 DR HAMAP; MF_00559; -; 1.
 DR InterPro; IPR000173; GAP_dhndrogenase.

Query Match 1.7%; Score 7; DB 1; Length 360;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 PGVVPGA 271
DB 49 PGVVPGA 55

```
RESULT 31
INCC2_ECOLI
ID INCC2_ECOLI STANDARD; PRT; 364 AA.
AC P07673;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Protein incC.
GN INCC.
OS Escherichia coli.
OG Plasmid IncP-alpha RK2.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86232584; PubMed=3520485;
RA Thomas C.M., Smith C.A.;
RT "The trfB region of broad host range plasmid RK2: the nucleotide
sequence reveals incC and key regulatory gene trfB/korB/kord as
overlapping genes.";
RL Nucleic Acids Res. 14:4453-4469 (1986).
RN [2]
RP SEQUENCE OF 330-364 FROM N.A.
RX MEDLINE=88015606; PubMed=3309894;
RA Theophilus B.D.B., Thomas C.M.;
RT "Nucleotide sequence of the transcriptional repressor gene korB which
plays a key role in regulation of the copy number of broad host range
plasmid RK2.";
RL Nucleic Acids Res. 15:7443-7450 (1987).
RN [3]
RP SEQUENCE OF 342-364 FROM N.A.
RX MEDLINE=88118923; PubMed=3430606;
RA Kornacki J.A., Balderes P.J., Figurski D.H.;
RT "Nucleotide sequence of korB, a replication control gene of broad
host-range plasmid RK2.";
RL J. Mol. Biol. 198:211-222 (1987).
CC -1- FUNCTION: THIS IS ONE OF THE PROTEINS ENCODED BY THE TRFB OPERON;
CC IT IS INVOLVED IN PLASMID MAINTENANCE AND REPLICATION.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative initiation;
CC Comment=2 isoforms, Long (shown here) and Short/Small, are
produced by alternative initiation;
CC -1- SIMILARITY: BELONGS TO THE PARA FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X03962; CAA27595.1; --
EMBL; Y00448; CAA68502.1; --
EMBL; X06543; CAA29789.1; --
DR IPR000707; ATPase_Para.
DR Pfam; PF00991; Para; 1.
KW Plasmid; DNA replication; Alternative initiation.
FT CHAIN 1 364 PROTEIN INCC, ISOFORM LONG.
FT CHAIN 106 364 PROTEIN INCC, ISOFORM SHORT.
FT INIT_MET 106 106 FOR ISOFORM SHORT.
SQ SEQUENCE 364 AA; 38134 MW; 36241B743C53E0BE CRC64;
```

Query Match 1.7%; Score 7; DB 1; Length 364;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 269 PGAGSGA 275
DB 19 PGAGSGA 25

```
RESULT 32
IA01_PONPY
ID IA01_PONPY STANDARD; PRT; 365 AA.
AC P16211;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Class I histocompatibility antigen, A-1 alpha chain precursor.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90201944; PubMed=1690682;
RA Lawlor D.A., Warren E., Ward F.E., Parham P.;
RT "Comparison of Class I MHC alleles in humans and apes.";
RL Immunol. Rev. 113:147-185 (1990).
CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
THE IMMUNE SYSTEM.
CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
MICROGLOBULIN).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; M30680; AAA88035.1; --
DR FIR; I84432; I84432.
DR HSP; Q95352; 1HKH.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR ProDom; PD000050; MHC_I; 1.
DR SMART; SM00407; IGcl_1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 365 CLASS I HISTOCOMPATIBILITY ANTIGEN,
A-1 ALPHA CHAIN.
FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.
FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.
FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.
FT DOMAIN 299 308 CONNECTING PEPTIDE.
FT TRANSMEM 309 332 CYTOPLASMIC TAIL.
FT DOMAIN 333 365 BY SIMILARITY.
FT DISULFID 125 188 BY SIMILARITY.
FT DISULFID 227 283 BY SIMILARITY.
FT CARBOHYD 110 110 N-LINKED (GLCNAC...) (BY SIMILARITY).
SQ SEQUENCE 365 AA; 40658 MW; 11A5BC183009CF70 CRC64;
```

Query Match 1.7%; Score 7; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 147 YSQASN 153
Db 344 YSQASN 350

RESULT 33
T2AI ACICA
ID T2AI ACICA STANDARD; PRT; 366 AA.
AC P24546;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Type II restriction enzyme AccI (EC 3.1.21.4) (Endonuclease AccI)
DE (R:AccI).
GN ACCIR.
OS Acinetobacter calcoaceticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=471;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-10.
RC STRAIN=ATCC 49823;
RX MEDLINE=91345839; PubMed=1368703;
RA Kawakami B., Hilzheber C., Nagatomo M., Oka M.;
RT "Cloning and nucleotide sequences of the AccI restriction-modification
RT genes in Acinetobacter calcoaceticus.";
RL Agric. Biol. Chem. 55:1553-1559(1991).
CC -!- FUNCTION: RECOGNIZES THE DOUBLE-STRANDED SEQUENCE GTMKAC AND
CC CLEAVES AFTER T-2.
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage of DNA to give
CC specific double-stranded fragments with terminal 5'-phosphates.
CC -!- SUBUNIT: Homotrimer.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; D10671; BRA01522.1; -.
CC F01; J00469; J00469.
CC REBASE; 18; AccI.
CC Restriction system; Hydrolase; Nuclease; Endonuclease.
CC SEQUENCE 366 AA; 42495 MW; DB663B74351C415E CRC64;

Query Match 1.7%; Score 7; DB 1; Length 366;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 110 ARTPTQA 116
Db 28 ARTPTQA 34

RESULT 34
SGK2 MOUSE
ID SGK2 MOUSE STANDARD; PRT; 367 AA.
AC Q9QZS5; Q9R0P6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Serine/threonine-protein kinase Sgk2 (EC 2.7.1.-)
DE (Serum/glucocorticoid regulated kinase 2).
GN SGK2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=20018032; PubMed=10548550;

```

```

RA Kobayashi T., Deak M., Morrice N., Cohen P.;
RT "Characterization of the structure and regulation of two novel
RL isoforms of serum- and glucocorticoid-induced protein kinase.";
RL Biochem. J. 344:189-197(1999).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Haegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hilli D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H.A., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazier K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Guetincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Kono A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempie C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Haehizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP -!- FUNCTION: Involved in the activation of potassium channels (By
RP similarity).
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- ALTERNATIVE PRODUCTS:
CC -!- Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9QZS5-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9QZS5-2; Sequence=VSP_004933;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -!- CAUTION: Not regulated by serum or glucocorticoids.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF169033; AAF12756.1; -.
CC EMBL; AK050009; BAC34031.1; -.
CC EMBL; BC026549; AAH26549.1; -.
CC HSSP; P05132; 1CTP.
CC MGI; 1351318; Sgk2.
CC InterPro; IPR000961; Pkinase C.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC Pfam; PF00433; pkinase C; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00133; S_TK_X; 1.
CC SMART; SM00220; S_TKC; 1.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS00111; PROTEIN KINASE DOM; 1.
CC PROSITE; PS00108; PROTEIN KINASE ST; 1.
CC Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Alternative splicing
KW DOMAIN 35 292
FT NP_BIND 41 49 ATP (BY SIMILARITY).
FT BINDING 64 64 ATP (BY SIMILARITY).
FT ACT_SITE 159 159 BY SIMILARITY.
FT MOD_RES 193 193 PHOSPHORYLATION (BY PDPK1)
FT (BY SIMILARITY)
FT VARSPIC 171 199 Missing (in isoform 2).
FT /FTID=VSP 004933.
FT CONFLICT 77 77 MISSING (IN REF. 3).
FT SEQUENCE 367 AA; 41359 MW; 668C04B1A1E9E33A CRC64;

Query Match 1.7%; Score 7; DB 1; Length 367;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 ASSSGAS 85
DB 343 ASSSGAS 349
|||||||

RESULT 35
ADHP_RABIT
ID ADHP_RABIT STANDARD; PRT; 378 AA.
AC O46649;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Alcohol dehydrogenase class II isozyme 1 (EC 1.1.1.1).
GN ADH2-1.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Liver;
RX MEDLINE=98151253; PubMed=9492289;
RA Svensson S., Hedberg J., Hoesog J.-O.;
RT "Structural and functional divergence of class II alcohol
RT dehydrogenase -- cloning and characterisation of rabbit liver isoforms
RT of the enzyme.";
RL Eur. J. Biochem. 251:236-243(1998).
CC -!- CATALYTIC ACTIVITY: An alcohol + NAD(+) = an aldehyde or ketone +
CC NADH.
CC -!- COFACTOR: Binds 2 zinc ions per subunit (By similarity).
CC -!- SUBUNIT: Homodimer (By similarity).

```

```

CC -----
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the zinc-containing alcohol dehydrogenase
CC family. Class-II subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ002388; CAA05362.1; -.
CC HSSP; P00325; 1DEH.
CC InterPro; IPR002328; ADH_zinc.
CC InterPro; IPR002085; Adh_zn_family.
CC Pfam; PF00107; ADH_zinc_N; 1.
CC PROSITE; PS00059; ADH_ZINC; 1.
KW Oxidoreductase; Zinc; Metal-binding; NAD; Multigene family.
FT INIT_MET 0 0
FT METAL 46 46 ZINC 1 (CATALYTIC) (BY SIMILARITY).
FT METAL 68 68 ZINC 1 (CATALYTIC) (BY SIMILARITY).
FT METAL 98 98 ZINC 2 (BY SIMILARITY).
FT METAL 101 101 ZINC 2 (BY SIMILARITY).
FT METAL 104 104 ZINC 2 (BY SIMILARITY).
FT METAL 112 112 ZINC 2 (BY SIMILARITY).
FT METAL 175 175 ZINC 1 (CATALYTIC) (BY SIMILARITY).
FT SEQUENCE 378 AA; 40105 MW; EA21E1EF5920FD CRC64;

Query Match 1.7%; Score 7; DB 1; Length 378;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 TLKGTFN 13
DB 318 TLKGTFN 324
|||||||

RESULT 36
SUCC_CHLTR
ID SUCC_CHLTR STANDARD; PRT; 386 AA.
AC O84828;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Succinyl-CoA synthetase beta chain (EC 6.2.1.5) (SCS-beta).
GN SUCC OR CT821.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/OW-3/Cx;
RX MEDLINE=9900809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis.";
RL Science 282:754-759(1998).
CC -!- CATALYTIC ACTIVITY: ATP + succinate + CoA = ADP + succinyl-CoA +
CC phosphate.
CC -!- PATHWAY: Tricarboxylic acid cycle.
CC -!- SUBUNIT: Composed of an alpha chain and a beta chain (By
CC similarity).
CC -!- SIMILARITY: Belongs to the succinate/malate CoA ligase beta
CC subunit family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```


RC STRAIN=129; TISSUE=Kidney;
RX MEDLINE=98392851; PubMed=9776250;
RA Frank S., Zoll B.;
RT "Mouse HNF-3/fork head homolog-1-like gene: structure, chromosomal
RT location, and expression in adult and embryonic kidney";
RL DNA Cell Biol. 17:679-688 (1998).
RN [2]
RP REVISIONS.
RA Pasche B., Bieller A., Zoll B.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP DISEASE.
RX MEDLINE=21207067; PubMed=11309849;
RA Hong H.-K., Noveroske J.K., Headon D.J., Liu T., Sy M.S.,
RA Justice M.J., Chakravarti A.;
RT "The winged helix/forkhead transcription factor Foxq1 regulates
RT differentiation of hair in statin mice";
RL Genesis 29:163-171 (2001).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- TISSUE SPECIFICITY: Expressed in kidney and stomach. Expression in
CC the outer medulla of the kidney and the transitional epithelium.
CC -1- DISEASE: Defects in FOXQ1 are the cause of the statin (sa)
CC phenotype and results in a silky coat with high sheen arising from
CC structurally abnormal medulla cells and defects in differentiation
CC of the hair shaft.
CC -1- SIMILARITY: Contains 1 fork-head domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AF010405; AAC12973.2; -.
DR HSSP; Q63245; 2HFH.
DR MGP; MGI:1298228; Foxq1.
DR InterPro; IPR001766; TF_Fork_head.
DR Pfam; PF00250; Fork_head; 1.
DR PRINTS; PR00053; FORKHEAD.
DR ProDom; PD000425; TF_Fork_head; 1.
DR SMART; SM00339; FH; 1.
DR PROSITE; PS00657; FORK HEAD 1; 1.
DR PROSITE; PS00658; FORK HEAD 2; 1.
DR PROSITE; PS00039; FORK HEAD 3; 1.
KW DNA-binding; Nuclear protein; Transcription regulation.
FT DNA BIND 115 210 FORK-HEAD.
SQ SEQUENCE 400 AA; 41342 MW; E6C3B5F53F3BC42B CRC64;

Query Match 1.7%; Score 7; DB 1; Length 400;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 PQAPAA 136
Dy 236 PQAPAA 242
|||||

RESULT 39
PROV_ECOLI
ID PROV_ECOLI STANDARD; PRT; 400 AA.
AC P14175;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glycine betaine/L-proline transport ATP-binding protein prov.
GN PROV OR B2677.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=89197759; PubMed=2649479;
RA Gowrishankar J.;
RT "Nucleotide sequence of the osmoregulatory proU operon of Escherichia
RT coli";
RL J. Bacteriol. 171:1923-1931 (1989).
RN [2]
RP ERRATUM.
RA Gowrishankar J.;
RL J. Bacteriol. 172:1165-1165 (1990).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MGI1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12";
RL Science 277:1453-1474 (1997).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=K12; PubMed=9205837;
RX MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Aiba H., Baba H., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
RA Oahina T., Oyama S., Saito N., Sampa G., Satoh Y., Sivasubram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT K-12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features";
RL DNA Res. 4:91-113 (1997).
RN [5]
RP SEQUENCE OF 1-224 FROM N.A.
RX STRAIN=K12;
RX MEDLINE=90113884; PubMed=2691838;
RA Stirling D.A., Hulton C.S.J., Waddell L., Park S.F., Stewart G.S.A.B.,
RA Booth I.R., Higgins C.F.;
RT "Molecular characterization of the proU loci of Salmonella
RT typhimurium and Escherichia coli encoding osmoregulated glycine
RT betaine transport systems";
RL Mol. Microbiol. 3:1025-1038 (1989).
CC -1- FUNCTION: INVOLVED IN A MULTICOMPONENT BINDING-PROTEIN-DEPENDENT
CC -1- TRANSPORT SYSTEM FOR GLYCINE BETAINES/L-PROLINE.
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
CC -1- SIMILARITY: Contains 2 CBS domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M24856; AAA24427.1; -.
DR EMBL; AE000352; AAC75724.1; -.
DR EMBL; D90891; BAA16542.1; -.
DR EMBL; X52694; CAA36923.1; -.
DR PIR; JS0128; BVECPV.
DR HSP; Q58663; IG5H.
DR ECO2DBASE; G033.5; 6TH EDITION.
DR EcoGene; EG10771; prov.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR000644; CBS domain.
DR InterPro; IPR005892; Prov.
DR Pfam; PF00005; ABC_tran; 1.
DR Pfam; PF00571; CBS; 1.
DR ProDom; PD000006; ABC_transporter; 1.

```

DR SMART; SM00382; AAA; 1.
DR SMART; SM00116; CBS; 1.
DR TIGRFAMs; TIGR01186; prov; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW Inner membrane; Amino-acid transport; Transport; ATP-binding; Repeat;
KW CBS domain; Complete proteome.
FT DOMAIN 29 265 ABC_TRANSPORTER.
FT DOMAIN 279 334 CBS 1.
FT DOMAIN 340 393 CBS 2.
FT NP BIND 61 68 ATP (BY SIMILARITY).
FT NP BIND 61 68 ATP (BY SIMILARITY).
SQ SEQUENCE 400 AA; 44162 MW; 45A98D45F028B1B9 CRC64;

Query Match
Best Local Similarity 1.7%; Score 7; DB 1; Length 400;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 275 ALDALRQ 281
DB 145 ALDALRQ 151

RESULT 40
PROV_SALTY STANDARD; PRT; 400 AA.
AC P17328;
DT 01-AUG-1990 (Rel. 15; Created)
DT 28-FEB-2003 (Rel. 41; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Glycine betaine/L-proline transport ATP-binding protein prov.
GN PROV OR STM2809.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=90113884; PubMed=2691838;
RA Stirling D.A., Hulton C.S.J., Waddell L., Park S.F., Stewart G.S.A.B.,
RA Booth I.R., Higgins C.F.;
RT "Molecular characterization of the proU loci of Salmonella
RT typhimurium and Escherichia coli encoding osmoregulated glycine
RT betaine transport systems."
RL Mol. Microbiol. 3:1025-1038(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856(2001).
RN [3]
RP SEQUENCE OF 1-291 FROM N.A.
RX MEDLINE=89359099; PubMed=2548994;
RA "Nucleotide sequence of the transcriptional control region of the
RT osmotically regulated proU operon of Salmonella typhimurium and
RT identification of the 5' endpoint of the proU mRNA."
RL J. Bacteriol. 171:4694-4706(1989).
RN [4]
RP SEQUENCE OF 1-196 FROM N.A.
RX MEDLINE=94253018; PubMed=8195103;
RA Jordan A., Gilbert I., Barbe J.;
RT "Cloning and sequencing of the genes from Salmonella typhimurium
RT encoding a new bacterial ribonucleotide reductase."
RL J. Bacteriol. 176:3420-3427(1994).
RN [5]
RP FUNCTION: INVOLVED IN A MULTICOMPONENT BINDING-PROTEIN-DEPENDENT
CC
```

```

CC TRANSPORT SYSTEM FOR GLYCINE BETAINE/L-PROLINE.
CC -! SUBCELLULAR LOCATION: Inner membrane-associated.
CC -! SIMILARITY: Belongs to the ABC transporter family.
CC -! SIMILARITY: Contains 2 CBS domains.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; X52693; CAA36921.1; -.
CC EMBL; AE008828; AAL21694.1; -.
CC EMBL; M26063; AAR8621.1; -.
CC EMBL; X73226; CAA51696.1; -.
CC PIR; S05374; QREBVT.
CC HSSP; Q58663; 1G6H.
CC StyGene; SG10312; prov.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR003439; ABC_transporter.
CC InterPro; IPR000644; CBS_domain.
CC InterPro; IPR005892; Prov.
CC Pfam; PF00005; ABC_tran; 1.
CC Pfam; PF00571; CBS; 2.
CC ProDom; PD000006; ABC_transporter; 1.
CC SMART; SM00382; AAA; 1.
CC SMART; SM00116; CBS; 1.
CC TIGRFAMs; TIGR01186; prov; 1.
CC PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
CC PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW Inner membrane; Amino-acid transport; Transport; ATP-binding; Repeat;
KW CBS domain; Complete proteome.
FT DOMAIN 29 265 ABC_TRANSPORTER.
FT DOMAIN 279 334 CBS 1.
FT NP BIND 61 68 ATP (BY SIMILARITY).
FT NP BIND 61 68 ATP (BY SIMILARITY).
FT CONFLICT 158 158 A -> R (IN REF. 1).
SQ SEQUENCE 400 AA; 44124 MW; FDF58DE0A9A792C2 CRC64;

Query Match
Best Local Similarity 1.7%; Score 7; DB 1; Length 400;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 275 ALDALRQ 281
DB 145 ALDALRQ 151

RESULT 41
SGK2_HUMAN STANDARD; PRT; 427 AA.
ID SGK2_HUMAN STANDARD; PRT; 427 AA.
AC Q9HBY6; Q9URK6;
DT 28-FEB-2003 (Rel. 41; Created)
DT 28-FEB-2003 (Rel. 41; Last sequence update)
DT 15-SEP-2003 (Rel. 42; Last annotation update)
DE Serine/threonine-protein kinase Sgk2 (EC 2.7.1.37)
DE (Serum/glucocorticoid regulated kinase 2).
GN SGK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1 AND 2), PHOSPHORYLATION ON THR-253, AND
RP MUTAGENESIS OF SER-416.
RX MEDLINE=20018032; PubMed=10548550;
RA Kobayashi T., Deak M., Morrice N., Cohen P.;
RT "Characterization of the structure and regulation of two novel
RT isoforms of serum- and glucocorticoid-induced protein kinase."
RL Biochem. J. 344:189-197(1999).
RN [2]
```

SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.V., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
 RA Ellington A.G., Griffiths J.A., Frazer A.A., French L., Garner P.,
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Lehaealaho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConnachie L.J., McMay K., McMurray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
 RA Stuke C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Colon;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marasina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McWay P.C., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.J., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Pahay J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP FUNCTION.
 RX MEDLINE=22284526; PubMed=12397388;
 RA Gamber N., Fillon S., Peng Y., Friedrich B., Lang P.A., Henke G.,
 RA Huber S.M., Kobayashi T., Cohen P., Lang P.;
 RT "K(+) channel activation by all three isoforms of serum- and
 glucocorticoid-dependent protein kinase SGK.";
 RL Pflugers Arch. 445:60-66(2002).
 CC -1- FUNCTION: Involved in the activation of potassium channels.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=2; Synonyms=beta;
 CC IsoId=Q9HBY8-1; Sequence=Displayed;
 CC Name=1; Synonyms=alpha;
 CC IsoId=Q9HBY8-2; Sequence=VSP_004932;
 CC -1- TISSUE SPECIFICITY: Highly expressed in liver, kidney and
 pancreas, and at lower levels in brain.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -1- CAUTION: Not regulated by serum or glucocorticoids.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL; AF166470; AAG17012.1; -;
 DR EMBL; AF169034; AAF12757.2; -;
 DR EMBL; Z98752; CAC18509.1; -;
 DR EMBL; BC014037; AAH14037.1; -;
 DR HSSP; P05132; 1ATP.
 DR Genew; HGNC:13900; SGK2.
 DR MIM; 607589; -;
 DR InterPro; IPR000961; Kinase C.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF00433; Pkinase_C; 1.
 DR ProDom; PD0000001; Prot_kinase; 1.
 DR SMART; SM00133; S_TK_X; 1.
 DR SMART; SM00220; S_TK; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW Phosphatase; Serine/threonine-protein kinase; ATP-binding;
 KW Phosphorylation; Alternative splicing.
 FT DOMAIN 95 352
 FT NP_BIND 101 109
 FT BINDING 124 124
 FT ACT_SITE 219 219
 FT MOD_RES 253 253
 FT VARSPPLIC 1 60
 FT FTID=VSP_004932.
 FT MUTAGEN 416 416 S->D: INCREASED ACTIVATION.
 SQ SEQUENCE 427 AA; 47604 MW; D8F0FA6DF54B1370 CRC64;
 Query Match 1.7%; Score 7; DB 1; Length 427;
 Best Local Similarity 100.0%; Pred. No. 96;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 79 ASSSGAS 85
 DB 403 ASSSGAS 409
 RESULT 42
 YAP1_CHICK
 ID_YAP1_CHICK STANDARD; PRT; 448 AA.
 AC P46936;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE 65 kDa Yes-associated protein (YAP65).
 GN YAP1 OR YAP65
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]_TaxID=9031;
 RP SEQUENCE FROM N.A.
 RC STRAIN=White leghorn;
 RX MEDLINE=94309887; PubMed=8035999;
 RA Sudol M.;
 RT "Yes-associated protein (YAP65) is a proline-rich phosphoprotein that
 RT binds to the SH3 domain of the Yes proto-oncogene product.";
 RL Oncogene 9:2145-2152(1994).
 CC -1- FUNCTION: BINDS TO THE SH3 DOMAIN OF THE YES KINASE. ALSO BINDS TO
 CC OTHER SIGNALING MOLECULES THAT CONTAIN SH3 DOMAINS INCLUDING NCK,
 CC CRK AND SRC.

```

CC -1- PTM: PHOSPHORYLATED ON SERINE RESIDUES.
CC -1- SIMILARITY: Contains 1 WW domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X76483; CAA54021.1; -.
DR PIR; I50730; I50730.
DR InterPro; IPR005153; MbCh.
DR InterPro; IPR002349; WW.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00397; WW; 1.
DR PRINTS; PR00403; WWDOMAIN.
DR SMART; SM00456; WW; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 1.
DR PROSITE; PS0020; WW_DOMAIN_2; 1.
KW Phosphorylation.
FT DOMAIN 169 202 WW
FT SEQUENCE 448 AA; 47822 MW; 719CC8D0F879A38D CRC64;
SQ
Query Match 1.7%; Score 7; DB 1; Length 448;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 269 PGAGSGA 275
DB 28 PGAGSGA 34
|||||
RESULT 43
ANT3 MOUSE
ID ANT3_MOUSE STANDARD; PRT; 465 AA.
AC P32261;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Antithrombin-III precursor (ATIII).
GN SERPINC1 OR AT3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93069082; PubMed=1440494;
RA Wu J.K., Sheffield W.P., Blajchman M.A.;
RT "Molecular cloning and cell-free expression of mouse antithrombin
RT III."
RL Thromb. Haemost. 68:291-296(1992).
CC -1- FUNCTION: MOST IMPORTANT SERINE PROTEASE INHIBITOR IN PLASMA
CC THAT REGULATES THE BLOOD COAGULATION CASCADE. AT-III INHIBITS
CC THROMBIN AS WELL AS FACTORS IXA, XA AND XIA. ITS INHIBITORY
CC ACTIVITY IS GREATLY ENHANCED IN THE PRESENCE OF HEPARIN.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: Plasma.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S47225; AAB23965.1; -.
DR HSSP; P01008; IATH.
DR MGD; MGI:88095; Serpinci.

```

```

DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Serpin; Serine protease inhibitor; Glycoprotein; Heparin-binding;
KW Plasma; Blood coagulation; Signal.
FT CHAIN 1 32 BY SIMILARITY.
FT BINDING 33 465 ANTITHROMBIN-III.
FT BINDING 82 82 HEPARIN (BY SIMILARITY).
FT BINDING 162 162 HEPARIN (BY SIMILARITY).
FT BINDING 178 178 HEPARIN (BY SIMILARITY).
FT ACT_SITE 426 427 REACTIVE BOND.
FT DISULFID 41 161 BY SIMILARITY.
FT DISULFID 54 128 BY SIMILARITY.
FT DISULFID 280 463 BY SIMILARITY.
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 168 168 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 225 225 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 465 AA; 52003 MW; 5CE087E98874E35D CRC64;
Query Match 1.7%; Score 7; DB 1; Length 465;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 269 PGAGSGA 275
DB 4 PGAGSGA 10
|||||
RESULT 44
FXD3 MOUSE
ID FXD3_MOUSE STANDARD; PRT; 465 AA.
AC G61050;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Forkhead box protein D3 (HNF3/FH transcription factor genesis)
DE DE
DE FOXD3 OR HPH2.
GN FOXD3 OR HPH2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Hromas R.A., Costa R.H., Xu D., Sutton J.L.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Labosky P.A., Kaestner K.H.;
RT "The winged helix transcription factor Hfh2 is expressed in neural
RT crest and spinal cord during mouse development."
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PROBABLE TRANSCRIPTION FACTOR.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Contains 1 fork-head domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U41047; AAA87569.1; -.
DR EMBL; AF067421; AAC28352.1; -.
DR HSSP; Q63245; 2HPH.
DR TRANSFAC; T04166; -.
DR MGD; MGI:1347473; Foxd3.
DR InterPro; IPR001766; TF_Fork_head.
DR Pfam; PF00250; Fork_head; 1.

```


use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR HAWAP; MF_00571; -; 1.
DR InterPro; IPR000766; Galp transf II.
DR InterPro; IPR005850; Galp Utransf C.
DR InterPro; IPR005849; Galp_Utransf_N.
DR InterPro; IPR005934; Galp2.
DR Pfam; PF02744; Galp_UDP_tr.C; 1.
DR Pfam; PF01087; Galp_UDP_transf; 1.
DR TIGRFAMs; TIGR01239; Galp2; 1.
DR PROSITE; PS01163; GAL P UDP TRANSF II; 1.
KW Transferase; Nucleotidyltransferase; Galactose metabolism.
SQ SEQUENCE 486 AA; 54161 MW; 91A8EDE4F449B035 CRC64;

Query Match 1.7%; Score 7; DB 1; Length 486;
Best Local Similarity 100.0%; Pred.No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 203 PENVEAQ 209
Db 444 PENVEAQ 450

RESULT 47

Y222_CABEL STANDARD; PRT; 498 AA.
ID Y222_CABEL
AC Q21339;
DT 01-NOV-1997 (Rel. 35, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein K08E5.2 in chromosome III.
GN K08E5.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Kershaw J.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Durbin R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE SLC13A FAMILY OF TRANSPORTERS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; Z30974; CAA83225.2; -;
DR WormPep; K08E5.2; CE28598.
DR InterPro; IPR001898; Na/sul_sympot.
DR Pfam; PF00939; Na_sulph_symp; 1.
DR TIGRFAMs; TIGR00785; Gass; 1.
DR PROSITE; PS01271; NA SULFATE; 1.
KW Hypothetical protein; Transmembrane; Transport.
FT TRANSMEM 24 44 POTENTIAL.
FT TRANSMEM 94 114 POTENTIAL.
FT TRANSMEM 151 171 POTENTIAL.
FT TRANSMEM 200 220 POTENTIAL.
FT TRANSMEM 261 281 POTENTIAL.

FT TRANSMEM 284 304 POTENTIAL.
FT TRANSMEM 332 352 POTENTIAL.
FT TRANSMEM 371 391 POTENTIAL.
FT TRANSMEM 428 448 POTENTIAL.
SQ SEQUENCE 498 AA; 54517 MW; FAEDE2FAA694BCB1 CRC64;

Query Match 1.7%; Score 7; DB 1; Length 498;
Best Local Similarity 100.0%; Pred.No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 264 NPGVVRG 270
Db 270 NPGVVRG 276

RESULT 48

DHAL_MOUSE STANDARD; PRT; 500 AA.
ID DHAL_MOUSE
AC P24549;
DT 01-MAR-1992 (Rel. 21, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Aldehyde dehydrogenase 1A1 (EC 1.2.1.3) (Aldehyde dehydrogenase, cytosolic) (ALDH class 1) (ALDH1) (ALDH-E1).
GN ALDH1A1 OR ALDH1 OR AHD2 OR AHD-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c, and C57BL/6; TISSUE=Liver;
RX MEDLINE=91276281; PubMed=2055490;
RA Rongoparut P., Weaver S.;
RT "Isolation and characterization of a cytosolic aldehyde dehydrogenase-encoding cDNA from mouse liver."
RL Gene 101:261-265(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c, C57BL/6J, and 129/REJ; TISSUE=Liver;
RX MEDLINE=9508515; PubMed=7993664;
RA Bond S.L., Singh S.M.;
RT "DNA sequence analysis of the cytosolic acetaldehyde dehydrogenase gene (Ahd-2) in mouse strains with variable ethanol preferences."
RL Biochem. Med. Metab. Biol. 52:155-159(1994).
CC -!- FUNCTION: IN ADDITION TO THE ACTIVITY ON ACETALDEHYDE AND RELATED SUBSTRATES, IS ALSO INVOLVED IN THE OXIDATION OF ALDEHYDES DERIVED FROM BIOGENIC AMINES SUCH AS EPINEPHRINE AND NOREPINEPHRINE, AS WELL AS THE ALDEHYDES GENERATED VIA LIPID PEROXIDATION. BINDS FREE RETINAL AND CELLULAR RETINOL-BINDING PROTEIN-BOUND RETINAL. CAN CONVERT/OXIDIZE RETINALDEHYDE TO RETINOIC ACID (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: An aldehyde + NAD(+) + H(2)O = an acid + NADH.
CC -!- PATHWAY: Ethanol utilization; second step.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE LIVER, LUNG, AND TESTES. IS APPARENTLY NOT EXPRESSED AT DETECTABLE LEVELS IN KIDNEY, STOMACH, OVARY, HEART, AND BRAIN.
CC -!- SIMILARITY: Belongs to the aldehyde dehydrogenase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; M74570; AAA37202.1; -;
DR EMBL; M74571; AAA37203.1; -;
DR EMBL; S75713; AAB32754.2; -;
DR EMBL; S77047; -; NOT ANNOTATED_CDS.
DR PIR; JQ1004; JQ1004.

```

DR HSP; P51977; 1BX5.
DR SWISS-2DPAGE; P24549; MOUSE.
DR MGD; MGI:1353450; Aldhl1.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR Pfam; PF00171; aldedh; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
KW Oxidoreductase; NAD; Acetylation.
FT INIT MET 0 0 BY SIMILARITY.
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
FT NP_BIND 245 250 NAD (ADP PART) (BY SIMILARITY).
FT ACT_SITE 268 268 BY SIMILARITY.
FT BINDING 299 304 BY SIMILARITY.
FT BINDING 299 304 ANTABUSE.
FT CONFLICT 7 7 A -> R (IN REF. 1).
FT CONFLICT 44 44 T -> S (IN REF. 2).
FT CONFLICT 50 50 H -> Q (IN REF. 2).
FT CONFLICT 86 86 R -> C (IN REF. 1).
FT CONFLICT 457 457 I -> M (IN REF. 2).
SQ SEQUENCE 500 AA; 54318 MW; 3E428154E7214B54 CRC64;

Query Match 1.7%; Score 7; DB 1; Length 500;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 SPAQPAV 233
DB 2 SPAQPAV 8

RESULT 49
DHAI RAT
ID DHAI RAT STANDARD; PRT; 500 AA.
AC P51647; O09184;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aldehyde dehydrogenase 1A1 (EC 1.2.1.3) (Aldehyde dehydrogenase,
DE cytosolic) (ALDH class 1) (Retinal dehydrogenase) (RALDH) (RALDH1).
GN ALDH1A1 OR ALDH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Kidney;
RX MEDLINE=96125208; PubMed=8543180;
RA Bhat P.V., Labrecque J., Boutin J.-M., Lacroix A., Yoshida A.;
RT "Cloning of a cDNA encoding rat aldehyde dehydrogenase with high
RT activity for retinal oxidation.";
RL Gene 166:303-306 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=97212790; PubMed=9059608;
RA Kathmann E.C., Lipsky J.J.;
RT "A preliminary report on the cloning of a constitutively expressed rat
RT liver cytosolic ALDH cDNA by PCR.";
RL Adv. Exp. Med. Biol. 414:69-72 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Napoli J.L., Penzes P., Wang X., Sperkova Z.;
RT "Cloning of a rat cDNA encoding retinal dehydrogenase isozyme type
RT I and its expression in E. coli.";
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 1-18; 79-90; 95-120; 204-224; 236-257 AND 393-437.
RC STRAIN=Sprague-Dawley; TISSUE=Kidney;
RX MEDLINE=95134257; PubMed=7832787;
RA Labrecque J., Dumas F., Lacroix A., Bhat P.V.;
RT "A novel isoenzyme of aldehyde dehydrogenase specifically involved in

```

```

RT the biosynthesis of 9-cis and all-trans retinoic acid.";
RL Biochem. J. 305:681-684 (1995).
CC - FUNCTION: IS CAPABLE OF CONVERTING 9-CIS AND ALL-TRANS RETINAL TO
CC CORRESPONDING RETINOIC ACID WITH HIGH EFFICIENCY, 9-CIS RETINAL
CC BEING 2-FOLD MORE ACTIVE THAN ALL-TRANS RETINAL.
CC - CATALYTIC ACTIVITY: An aldehyde + NAD(+) + H(2)O = an acid + NADH.
CC - ENZYME REGULATION: INHIBITED BY CHLORAL HYDRATE.
CC - PATHWAY: Ethanol utilization; second step.
CC - SUBUNIT: Homotetramer (By similarity).
CC - SUBCELLULAR LOCATION: Cytoplasmic.
CC - TISSUE SPECIFICITY: STRONGLY EXPRESSED IN KIDNEY, LUNG, TESTIS,
CC INTESTINE, STOMACH, AND TRACHEA, BUT WEAKLY IN THE LIVER.
CC - SIMILARITY: Belongs to the aldehyde dehydrogenase family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; L42009; AAA96657.1; -.
DR EMBL; AF001896; AAC53304.1; -.
DR EMBL; AF001898; AAC53306.1; -.
DR EMBL; AF001897; AAC53305.1; -.
DR EMBL; U79118; AAB63423.1; -.
DR EIR; JC4524; JC4524.
DR PIR; JC5553; JC5553.
DR HSP; P51977; 1BX5.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR Pfam; PF00171; aldedh; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
KW Oxidoreductase; NAD; Acetylation.
FT INIT MET 0 0
FT MOD_RES 1 1 ACETYLATION (POTENTIAL).
FT NP_BIND 245 250 NAD (ADP PART) (BY SIMILARITY).
FT ACT_SITE 268 268 BY SIMILARITY.
FT ACT_SITE 302 302 BY SIMILARITY.
FT CONFLICT 99 99 R -> C (IN REF. 1).
FT CONFLICT 105 105 I -> M (IN REF. 4).
FT CONFLICT 169 169 N -> E (IN REF. 1).
SQ SEQUENCE 500 AA; 54327 MW; 93614C21A94B430D CRC64;

Query Match 1.7%; Score 7; DB 1; Length 500;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 SPAQPAV 233
DB 2 SPAQPAV 8

RESULT 50
DHAI RAT
ID DHAI RAT STANDARD; PRT; 500 AA.
AC P13601;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aldehyde dehydrogenase, cytosolic 1 (EC 1.2.1.3) (ALDH class 1)
DE (ALDH1) (ALDH-E1).
GN ALDH1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Kidney;
RX MEDLINE=89327272; PubMed=2753900;
RA Dunn T.J., Koleske A.J., Lindahl R., Pitot H.C.;
RT "Phenobarbital-inducible aldehyde dehydrogenase in the rat. cDNA

```

```

RT sequence and regulation of the mRNA by phenobarbital in responsive
RL rats." J. Biol. Chem. 264:13057-13065 (1989).
CC -1- FUNCTION: BINDS FREE RETINAL AND CELLULAR RETINOL-BINDING PROTEIN-
CC BOUND RETINAL. CAN CONVERT/OXIDIZE RETINALDEHYDE TO RETINOIC ACID
CC (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: An aldehyde + NAD(+) + H(2)O = an acid + NADH.
CC -1- PATHWAY: Ethanol utilization; second step.
CC -1- SUBUNIT: Homotetramer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- INDUCTION: By phenobarbital.
CC -1- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M23995; AAA40718.1; -.
CC DR PIR; A32616; A32616.
CC DR HSP; P51977; 1BX5.
CC DR InterPro; IPR002086; Aldehyde_dehydr.
CC DR Pfam; PF00171; aldehyd; 1.
CC DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
CC DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
CC KW Oxidoreductase; NAD.
CC FT INIT MET 0 BY SIMILARITY.
CC FT NP_BIND 245 250 NAD (ADP PART) (BY SIMILARITY).
CC FT ACT_SITE 268 268 BY SIMILARITY.
CC FT ACT_SITE 302 302 BY SIMILARITY.
CC SQ SEQUENCE 500 AA; 54428 MW; D6806A1AE29A0D56 CRC64;

Query Match 1.7%; Score 7; DB 1; Length 500;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 227 SPAQPAV 233
Db 2 SPAQPAV 8
|||||

RESULT 51
ID IRX3 MOUSE STANDARD; PRT; 507 AA.
AC P81067.
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Iroquois-class homeodomain protein IRX-3 (Iroquois homeobox protein
DE 3).
DN IRX3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A. PubMed=9486539;
RX MEDLINE=98145944; PubMed=9486539;
RA Boase A., Zulch A., Becker M.B., Torres M., Gomez-Skarmeta J.-L.,
RA Modolell J., Gruss P.;
RA "Identification of the vertebrate Iroquois homeobox gene family with
RT overlapping expression during early development of the nervous
RT system.";
RL Mech. Dev. 69:169-181(1997).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: BELONGS TO THE TALE/IRO HOMEBOX FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y15001; CAA75233.1; -.
CC DR HSP; P41778; 1DU6.
CC DR TRANSFAC; T02439; -.
CC DR MGD; MGI:1197522; Irx3.
CC DR InterPro; IPR001356; Homeobox.
CC DR InterPro; IPR003893; Iroquois_homeo.
CC DR Pfam; PF00046; homeobox; 1.
CC DR ProDom; PD000010; Homeobox; 1.
CC DR SMART; SM00389; HOX; 1.
CC DR PROSITE; PS00027; HOMEBOX_1; 1.
CC DR PROSITE; PS00071; HOMEBOX_2; 1.
CC KW Homeobox; DNA-binding; Nuclear protein.
CC FT DNA_BIND 130 192 HOMEBOX (TALE-TYPE).
CC FT DOMAIN 24 29 POLY-GLY.
CC FT DOMAIN 64 73 POLY-ALA.
CC FT DOMAIN 228 235 POLY-GLU.
CC FT DOMAIN 250 253 POLY-ARG.
CC FT DOMAIN 315 319 POLY-PRO.
CC FT DOMAIN 391 399 POLY-ALA.
CC FT DOMAIN 451 458 POLY-ALA.
CC SQ SEQUENCE 507 AA; 52694 MW; C785AFER83E22FF CRC64;

Query Match 1.7%; Score 7; DB 1; Length 507;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 210 PVARAPA 216
Db 319 PVARAPA 325
|||||

RESULT 52
ID NQOD PARDE STANDARD; PRT; 513 AA.
AC P29925.
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NADH-quinone oxidoreductase chain 13 (EC 1.6.99.5) (NADH dehydrogenase
DE I. chain 13) (NDH-1, chain 13).
DN NQO13.
OS Paracoccus denitrificans.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Paracoccus.
OX NCBI_TaxID=266;
[1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 13543;
RX MEDLINE=93136200; PubMed=8422400;
RA Xu X., Matsuno-Yagi A., Yagi T.;
RA "DNA sequencing of the seven remaining structural genes of the gene
RT cluster encoding the energy-transducing NADH-quinone oxidoreductase
RT of Paracoccus denitrificans.";
RL Biochemistry 32:968-981(1993).
CC -1- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
CC sulfur (Fe-S) centers, to quinones in the respiratory chain. The
CC immediate electron acceptor for the enzyme in this species is
CC believed to be ubiquinone. Couples the redox reaction to proton
CC translocation (for every two electrons transferred, four hydrogen
CC ions are translocated across the cytoplasmic membrane), and thus
CC conserves the redox energy in a proton gradient.
CC -1- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
CC -1- SUBUNIT: COMPOSED OF 14 DIFFERENT SUBUNITS. SUBUNITS NQO7-14
CC CONSTITUTE THE MEMBRANE SECTOR OF THE COMPLEX.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: TO POLYPEPTIDE 4 OF THE NADH-UBIQUINOL OXIDOREDUCTASE
CC OF CHLOROPLASTS OR MITOCHONDRIA.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L02334; AAA25599.1; -
CC PIR; I45456; I45456.
CC InterPro; IPR003918; NADHub oxford4.
CC InterPro; IPR001750; Oxidored_g1.
CC Pfam; PF00361; Oxidored_g1; 1.
CC PRINTS; PR01437; NUOXDRDTASE4.
CC OXidoreductase; NAD; Quinone; Ubiquinone; Transmembrane.
FT TRANSMEM 3 23 POTENTIAL.
FT TRANSMEM 34 54 POTENTIAL.
FT TRANSMEM 81 101 POTENTIAL.
FT TRANSMEM 112 132 POTENTIAL.
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 164 184 POTENTIAL.
FT TRANSMEM 211 231 POTENTIAL.
FT TRANSMEM 250 270 POTENTIAL.
FT TRANSMEM 277 297 POTENTIAL.
FT TRANSMEM 312 332 POTENTIAL.
FT TRANSMEM 340 360 POTENTIAL.
FT TRANSMEM 383 403 POTENTIAL.
FT TRANSMEM 418 438 POTENTIAL.
FT TRANSMEM 463 483 POTENTIAL.
SQ SEQUENCE 513 AA; 56417 MW; 8164DEACA20FE739 CRC64;

Query Match 1.7%; Score 7; DB 1; Length 513;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 OPAAPVA 104
Db 501 OPAAPVA 507
|||||

RESULT 53
THS3_HALVO
ID THS3_HALVO STANDARD; PRT; 524 AA.
AC Q9HHA2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Thermosome subunit 3 (Heat shock protein CCT3).
GN CCT3.
OS Halobacterium volcanii (Haloflex volcanii).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Haloflex.
OX NCBI_TaxID=2246;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DS70;
RA Kovacs E., Lund P.A.;
RT "Sequence of the cct3 gene, a third chaperonin homolog in the
RT halophilic archaeon Haloflex volcanii.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MOLECULAR CHAPERONE; BINDS UNFOLDED POLYPEPTIDES IN
CC VITRO, AND HAS A WEAK ATPASE ACTIVITY (BY SIMILARITY).
CC -1- SUBUNIT: FORMS AN OLIGOMERIC COMPLEX OF EIGHT-MEMBERED RINGS
CC (BY SIMILARITY).
CC -1- INDUCTION: By heat shock.
CC -1- SIMILARITY: Belongs to the TCP-1 chaperonin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF298660; AAG17906.1; -
CC HSSP; P48424; I46D.
CC InterPro; IPR002194; Chaperonin TCP-1.
CC InterPro; IPR001844; Chaperonin Cpn60.
CC InterPro; IPR002423; Cpn60/TCP-1.
CC Pfam; PF00118; Cpn60_TCP1; 1.
CC PRINTS; PR00298; CHAPERONIN60.
CC PRINTS; PR00304; TCOMPLEXTCP1.
CC PROSITE; PS00750; TCP1_1; 1.
CC PROSITE; PS00751; TCP1_2; 1.
CC PROSITE; PS00995; TCP1_3; 1.
KW Chaperone; ATP-binding; Heat shock.
SQ SEQUENCE 524 AA; 55261 MW; CBE1E2EAF91181E9 CRC64;

Query Match 1.7%; Score 7; DB 1; Length 524;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 266 GVPFGAG 272
Db 404 GVPFGAG 410
|||||

RESULT 54
PTB_MOUSE
ID PTB_MOUSE STANDARD; PRT; 527 AA.
AC P17225;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Polypyrimidine tract-binding protein 1 (PTB) (Heterogeneous nuclear
DE ribonucleoprotein I) (hnRNP I).
GN PTBP1 OR PTB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 312-341.
RX MEDLINE=92105132; PubMed=1722210;
RA Bothwell A.L.M., Ballard D.W., Philbrick W.M., Lindwall G.,
RA Maher S.E., Bridgett M.M., Jamison S.F., Garcia-Bianco M.A.;
RT "Murine polypyrimidine tract binding protein. Purification, cloning,
RT and mapping of the RNA binding domain.";
RL J. Biol. Chem. 266:24657-24663(1991).
CC -1- FUNCTION: PLAYS A ROLE IN PRE-MRNA SPLICING. BINDS TO THE
CC POLYPYRIMIDINE TRACT OF INTRONS. MAY PROMOTE THE BINDING OF U2
CC SNRNP TO PRE-MRNA.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DOMAIN: THE C-TERMINAL 195 AMINO ACIDS OF PTB ARE SUFFICIENT FOR
CC SPECIFIC RNA BINDING.
CC -1- SIMILARITY: Contains 4 RNA recognition motif (RRM) domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X52101; CAA36321.1; -
CC MGD; MGI:97791; Ptbp1.
CC GO; GO:0005634; C:nucleus; IDA.
CC InterPro; IPR006536; hnRNP-L_PTB.
CC InterPro; IPR000504; RNA_rec_mot.
CC Pfam; PF00076; rtm; 4.
CC SMART; SM00360; RRM; 3.
CC TIGRFAMs; TIGR01649; hnRNP-L_PTB; 1.
CC PROSITE; PS50102; RRM; 4.

```

```

DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
KW Nuclear protein; RNA-binding; mRNA splicing; Repeat.
FT DOMAIN 58 142 RNA-BINDING (RRM) 1.
FT DOMAIN 183 259 RNA-BINDING (RRM) 2.
FT DOMAIN 335 386 RNA-BINDING (RRM) 3.
FT DOMAIN 450 525 RNA-BINDING (RRM) 4.
FT DOMAIN 315 321 POLY-ALA.
SQ SEQUENCE 527 AA; 56478 MW; F18FDF376010D76A CRC64;

Query Match
Best Local Similarity 1.7%; Score 7; DB 1; Length 527;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 284 QFQALLQ 290
Db 220 QFQALLQ 226
|||||

RESULT 55
PTB_HUMAN
ID PTB_HUMAN STANDARD; PRT; 531 AA.
AC P26599;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Polypyrimidine tract-binding protein 1 (PTB) (Heterogeneous nuclear
DE ribonucleoprotein I) (hnRNP I) (57 kDa RNA-binding protein PTB-1).
GN PTB1 OR PTB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A., AND SEQUENCE OF 123-130 AND 219-238.
RP TISSUE=Placenta;
RX MEDLINE=91293583; PubMed=1906035;
RA Gil A., Sharp P.A., Jamison S.P., Garcia-Blanco M.A.;
RT "Characterization of cDNAs encoding the polypyrimidine tract-binding
RT protein.";
RL Genes Dev. 5:1224-1236 (1991).
[2]
SEQUENCE FROM N.A.
RX MEDLINE=91293584; PubMed=1906036;
RA Patton J.G., Mayer S.A., Tempst P., Nadal-Ginard B.;
RT "Characterization and molecular cloning of polypyrimidine
RT tract-binding protein: a component of a complex necessary for
RT pre-mRNA splicing.";
RL Genes Dev. 5:1237-1251 (1991).
[3]
SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92350668; PubMed=1641332;
RA Ghetti A., Pinol-Roma S., Michael W., Morandi C., Dreyfuss G.;
RT "hnRNP I, the polypyrimidine tract-binding protein: distinct nuclear
RT localization and association with hnRNAs.";
RL Nucleic Acids Res. 20:3671-3678 (1992).
[4]
SEQUENCE FROM N.A.
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,
RA Burkhart-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Garnes J.,
RA Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,
RA Attix C., Andreise T., Trankheim M., Amico-Keller G., Cosfield J.,
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,
RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
RA Olsen A.S., Carrano A.V.;
RT "Sequence analysis of a 3.5 Mb contig in human 19p13.3 containing a
RT serine protease gene cluster.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
[5]
SEQUENCE OF 123-129; 399-405; 411-426 AND 429-437.
RC TISSUE=Keratinocytes;
RX MEDLINE=93162043; PubMed=1286667;

```

```

RA Raamussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,
RA Vandeckerckhove J.;
RT "Microsequences of 145 proteins recorded in the two-dimensional gel
RT protein database of normal human epidermal keratinocytes.";
RL Electrophoresis 13:960-969 (1992).
CC -1- FUNCTION: PLAYS A ROLE IN PRE-MRNA SPLICING. BINDS TO THE
CC POLYPYRIMIDINE TRACT OF INTRONS. MAY PROMOTE THE BINDING OF U2
CC SNRNP TO PRE-MRNA.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=1;
CC IsoId=P26599-1; Sequence=Displayed;
CC Name=2; Synonyms=PTB2;
CC IsoId=P26599-2; Sequence=VSP_005802;
CC -1- SIMILARITY: Contains 4 RNA recognition motif (RRM) domains.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
-----
DR EMBL; X62006; CAA43973.1; -
DR EMBL; X65371; CAA46443.1; -
DR EMBL; X65372; CAA46444.1; -
DR EMBL; X60648; CAA43056.1; -
DR EMBL; X66975; CAA47386.1; -
DR EMBL; AC006273; AAC99798.1; -
DR PIR; S23016; S23016.
DR PDB; 1QM9; 03-JUL-00.
DR Aarnus/Ghent-2DFAGE; 1505; NEPHGE.
DR Genew; HGNC:9583; PTBP1.
DR GK; P26599; -
DR MIM; 60693; -
DR GO; GO:0005730; C:nucleolus; TAS.
DR GO; GO:0005654; C:nucleoplasm; TAS.
DR GO; GO:0008436; F:heterogeneous nuclear ribonucleoprotein; TAS.
DR GO; GO:0008187; F:poly-pyrimidine tract binding activity; TAS.
DR GO; GO:0006371; P:mRNA splicing; TAS.
DR InterPro; IPR006536; hRNP-L_PTB.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 4.
DR SMART; SM00360; RRM; 4.
DR TIGRFAMs; TIGR01649; hRNP-L_PTB; 1.
DR PROSITE; PS0102; RRM; 4.
DR PROSITE; PS0030; RRM_RNP_1; FALSE NEG.
KW Nuclear protein; RNA-binding; mRNA splicing; Repeat;
KW Alternative splicing; 3D-structure.
FT DOMAIN 59 143 RNA-BINDING (RRM) 1.
FT DOMAIN 137 260 RNA-BINDING (RRM) 2.
FT DOMAIN 384 411 RNA-BINDING (RRM) 3.
FT DOMAIN 454 529 RNA-BINDING (RRM) 4.
FT DOMAIN 316 323 POLY-ALA.
FT VARSPLIC 297 297 F->FASPVAGAGPPTFAIPQAA (in isoform 2).
SQ SEQUENCE 531 AA; 57221 MW; BE12D5EA21537BED CRC64;
/FTID=VSP_005802.

Query Match
Best Local Similarity 1.7%; Score 7; DB 1; Length 531;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 284 QFQALLQ 290
Db 221 QFQALLQ 227
|||||

RESULT 56
GAG_BAEVM
ID _GAG_BAEVM STANDARD; PRT; 537 AA.

```

P03341; P10268;
 21-JUL-1986 (Rel. 01, Created)
 21-MAR-1989 (Rel. 10, Last sequence update)
 28-FEB-2003 (Rel. 41, Last annotation update)
 GAG polyprotein [contains: inner coat protein P12; Core protein P15;
 Core shell protein P30; Nucleoprotein P10].
 GAG.
 OS Baboon endogenous virus (strain M7).
 OC Viruses; Retroviral viruses; Retroviridae; Mammalian type C retroviruses.
 OX NCBI_TaxID=11764;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kato S., Matsuo K., Nishimura N., Takahashi N., Takano T.;
 RT "The entire nucleotide sequence of baboon endogenous virus DNA: a
 RT chimeric genome structure of murine type C and simian type D
 RT retroviruses.";
 RL Jpn. J. Genet. 62:127-137(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83241915; PubMed=6408267;
 RA Tamura T.;
 RT "Provirus of M7 baboon endogenous virus: nucleotide sequence of the
 RT gag-pol region.";
 RL J. Virol. 47:137-145(1983).
 CC -!- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -!- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-POL
 CC POLYPEPTIDE.
 CC -!- SIMILARITY: Contains 1 CCHC-type zinc finger.
 CC -!- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN EXTENSIVELY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D10032; BAA00923.1; -;
 DR EMBL; X05470; CAA29027.1; -;
 DR EMBL; J02034; -; NOT ANNOTATED_CDS.
 DR PIR; J02060; FOMVM7.
 DR InterPro; IPR000840; Gag_MA.
 DR InterPro; IPR003036; Gag_P30.
 DR InterPro; IPR001878; ZnF_CCHC.
 DR Pfam; PF01140; Gag_MA; 1.
 DR Pfam; PF02093; Gag_P30; 1.
 DR Pfam; PF00098; zf-CCHC; 1.
 DR SMART; SM00343; ZnF_C2HC; 1.
 DR PROSITE; PS0158; ZF_CCHC; 1.
 DR Core protein; Coat protein; Nucleoprotein; Polyprotein; Myristate;
 KW Zinc-finger.
 FT CHAIN 2 112 INNER COAT PROTEIN P12.
 FT CHAIN 113 226 CORE PROTEIN P15.
 FT CHAIN 227 477 CORE SHELL PROTEIN P30.
 FT CHAIN 478 537 NUCLEOPROTEIN P10.
 FT ZN FING 501 518 CCHC-TYPE.
 FT LIPID 2 2 MYRISTATE.
 SQ SEQUENCE 537 AA; 60623 MW; 5DEEE4437CFCFB79 CRC64;
 Query Match 1.7%; Score 7; DB 1; Length 537;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 355 TPPEERA 361
 DB 446 TPPEERA 452
 RESULT 57
 SYK_RHILO
 ID SYK_RHILO STANDARD; PRT; 548 AA.
 AC Q98EC8;

28-FEB-2003 (Rel. 41, Created)
 28-FEB-2003 (Rel. 41, Last sequence update)
 28-FEB-2003 (Rel. 41, Last annotation update)
 Lysyl-tRNA synthetase [EC 6.1.1.6] (Lysine--tRNA ligase) (LysRS).
 GN LYS OR ML5534.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuura A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 CC -!- CATALYTIC ACTIVITY: ATP + L-lysine + tRNA(Lys) = AMP + diphosphate
 CC + L-lysyl-tRNA(Lys).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF003007; BAB52044.1; -;
 DR HAMAP; MF_00177; -; 1.
 DR InterPro; IPR002904; Lys tRNA-synt_1c.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR Pfam; PF01921; tRNA-synt_1f; 1.
 DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 FT SITE 52 60 "HIGH" REGION.
 FT SITE 300 304 "KMSK" REGION.
 FT BINDING 303 303 ATP (BY SIMILARITY).
 SQ SEQUENCE 548 AA; 61014 MW; BSB44B8DF3D89C29 CRC64;
 Query Match 1.7%; Score 7; DB 1; Length 548;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 331 SPEGGPG 337
 DB 495 SPEGGPG 501
 RESULT 58
 SYK_BRUME
 ID SYK_BRUME STANDARD; PRT; 551 AA.
 AC Q8YCM8;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Lysyl-tRNA synthetase [EC 6.1.1.6] (Lysine--tRNA ligase) (LysRS).
 GN LYS OR BME110500.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29459;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / Biotype 1;
 RX MEDLINE=20020109; PubMed=11756686;

RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mijer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
RA Jablonki L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyrides N., Overbeek R.,
RA "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448 (2002).
CC -1- CATALYTIC ACTIVITY: ATP + L-lysine + tRNA(Lys) = AMP + diphosphate
CC + L-lysyl-tRNA(Lys).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AC009686; AAC53742.1; -;
CC PIR; AC3572; AC3572.
CC HANAP; MF 00177; -; 1.
CC InterPro; IPR002904; Lys-trna-synt_1c.
CC InterPro; IPR001412; trna-synt_1.
CC Pfam; PF01921; trna-synt_1f; 1.
CC PROSITE; PS00178; AA TRNA_LIGASE I; 1.
CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Complete proteome.
CC SITE 54 62 "HIGH" REGION.
CC SITE 303 307 "KMSKS" REGION.
CC BINDING 306 306 ATP (BY SIMILARITY).
CC SEQUENCE 551 AA; 61699 MW; FC55DB8B469E1AC0 CRC64;
CC
CC Query Match 1.7%; Score 7; DB 1; Length 551;
CC Best Local Similarity 100.0%; Pred. No. 1.2e+02;
CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 331 SPEGGPG 337
CC DB 498 SPEGGPG 504
CC
CC RESULT 59
CC SYK BRUSU
CC ID SYK BRUSU STANDARD; PRT; 551 AA.
CC AC PS9225;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DE Lysyl-tRNA synthetase (EC 6.1.1.6) (Lysine--tRNA ligase) (LysRS).
CC GN LYSS OR BRA0790.
CC OS Brucella suis.
CC OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
CC OC Brucellaceae; Brucella.
CC OX NCBI_TaxID=29461;
CC [1]
CC PP SEQUENCE FROM N.A.
CC RC STRAIN=1330 / Biovar 1;
CC RX MEDLINE=22247741; PubMed=12271122;
CC RA Paulsen I.T., Seshadri R., Nelson K.E., Brinkac L.M., Beanan M.J.,
CC Read T.D., Dodson R.J., Unayam L., Durkin A.S., Kolonay J.F., Madupu R.,
CC Daugherty S.C., Deboy R.T., Shetty J., Malek J., Van Aken S.E.,
CC Nelson W.C., Ayodeji B., Kraul M., Shetty J., White O., Salzberg S.L.,
CC Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
CC Hoover D.L., Lindler L.E., Helling S.M., Boyle S.M., Fraser C.M.;
CC "The Brucella suis genome reveals fundamental similarities between
CC animal and plant pathogens and symbionts."
CC Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153 (2002).
CC -1- CATALYTIC ACTIVITY: ATP + L-lysine + tRNA(Lys) = AMP + diphosphate
CC + L-lysyl-tRNA(Lys).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

CC -1- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE014573; AAN33969.1; -;
CC TIGR; BRA0790; -; 1.
CC HANAP; MF 00177; -; 1.
CC PROSITE; PS00178; AA TRNA_LIGASE I; 1.
CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Complete proteome.
CC SITE 54 62 "HIGH" REGION.
CC SITE 303 307 "KMSKS" REGION.
CC BINDING 306 306 ATP (BY SIMILARITY).
CC SEQUENCE 551 AA; 61668 MW; 5E9B3CAB71A0EB80 CRC64;
CC
CC Query Match 1.7%; Score 7; DB 1; Length 551;
CC Best Local Similarity 100.0%; Pred. No. 1.2e+02;
CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 331 SPEGGPG 337
CC DB 498 SPEGGPG 504
CC
CC RESULT 60
CC ID PTB RAT STANDARD; PRT; 555 AA.
CC AC Q00438; Q63568;
CC DT 01-APR-1993 (Rel. 25, Created)
CC DT 01-APR-1993 (Rel. 25, Last sequence update)
CC DE Polypyrimidine tract-binding protein 1 (PTB) (Heterogeneous nuclear
CC DE ribonucleoprotein I) (hnRNP I) (Pyrimidine-binding protein) (PYBP).
CC GN PTB1 OR PTB OR PYBP OR TBF1.
CC OS Rattus norvegicus (Rat).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC OX NCBI_TaxID=10116;
CC [1]
CC PP SEQUENCE FROM N.A., AND SEQUENCE OF 185-205; 348-365 AND 517-549.
CC RC TISSUE=Hepatoma;
CC RX MEDLINE=92020211; PubMed=1681508;
CC RA Brunel F., Alzari P.M., Ferrara P., Zakari M.M.;
CC "Cloning and sequencing of PYBP, a pyrimidine-rich specific single
CC RT strand DNA-binding protein."
CC RL Nucleic Acids Res. 19:5237-5245 (1991).
CC [2]
CC PP SEQUENCE FROM N.A.
CC RA Sengupta P.;
CC RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PLAYS A ROLE IN PRE-MRNA SPLICING. BINDS TO THE
CC POLYPYRIMIDINE TRACT OF INTRONS. MAY PROMOTE THE BINDING OF U2
CC SNRNP TO PRE-MRNA.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Names=PYBP2;
CC IsoIdc=Q00438-1; Sequence=Displayed;
CC Names=PYBP1;
CC IsoIdc=Q00438-2; Sequence=VSP_005803;
CC -1- SIMILARITY: Contains 4 RNA recognition motif (RRM) domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)

DR EMBL; X60789; CAA43202.1; -;
 DR EMBL; X60790; CAA43203.1; -;
 DR EMBL; X74565; CAA52653.1; -;
 DR PIR; S15522; S15552.
 DR PIR; S36629; S36629.
 DR InterPro; IPR000536; HnRNP-L_PTB.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00076; rrm; 4.
 DR SMART; SM00360; RRM; 4.
 DR TIGRFAMs; TIGR01649; hnRNP-L_PTB; 1.
 DR PROSITE; PS0102; RRM; 4.
 DR PROSITE; PS0030; RRM_RNP_1; FALSE NEG.
 KW Nuclear protein; RNA-binding; mRNA splicing; Repeat;
 KW Alternative splicing.
 FT DOMAIN 58 142 RNA-BINDING (RRM) 1.
 FT DOMAIN 183 259 RNA-BINDING (RRM) 2.
 FT DOMAIN 361 435 RNA-BINDING (RRM) 3.
 FT DOMAIN 478 553 RNA-BINDING (RRM) 4.
 FT DOMAIN 340 347 POLY-ALA.
 FT VARSPLIC 297 321 Missing (in isoform PYBP1).
 FT CONFLICT 57 57 /FTID=VSP 005803.
 FT CONFLICT 164 164 S -> T (IN REF. 2).
 FT CONFLICT 311 320 A -> R (IN REF. 2).
 FT CONFLICT 542 549 VPSHLCHPSR -> GFPTFAIPAQA (IN REF. 2).
 FT CONFLICT 542 549 ENHHLRVS -> LGDNHHRK (IN REF. 1; AA
 SEQUENCE).
 SQ SEQUENCE 555 AA; 59353 MW; E73F22B54467117F CRC64;

Query Match 1.7%; Score 7; DB 1; Length 555;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 284 QFOALLQ 290
 DB 220 QFOALLQ 226

RESULT 61

PTB_PIG STANDARD; PRT; 557 AA.
 AC Q29099;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE Polypyrimidine tract-binding protein 1 (PTB) (Heterogeneous nuclear
 ribonucleoprotein I) (hnRNP I).
 GN PTBP1 OR PTB.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=96249475; PubMed=8654585;
 RA Niepmann M.;
 RT "Portine polypyrimidine tract-binding protein stimulates translation
 initiation at the internal ribosome entry site of foot-and-mouth-
 disease virus.";
 RL FEBS Lett. 388:39-42(1996).
 CC -1- FUNCTION: PLAYS A ROLE IN PRE-MRNA SPLICING. BINDS TO THE
 POLYPYRIMIDINE TRACT OF INTRONS. MAY PROMOTE THE BINDING OF U2
 SNRNP TO PRE-MRNA.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Contains 4 RNA recognition motif (RRM) domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)

DR EMBL; X93009; CAA63597.1; -;
 DR PIR; S68857; S68857.
 DR InterPro; IPR006536; HnRNP-L_PTB.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00076; rrm; 4.
 DR SMART; SM00360; RRM; 4.
 DR TIGRFAMs; TIGR01649; hnRNP-L_PTB; 1.
 DR PROSITE; PS0102; RRM; 4.
 DR PROSITE; PS0030; RRM_RNP_1; FALSE NEG.
 KW Nuclear protein; RNA-binding; mRNA splicing; Repeat.
 FT DOMAIN 59 143 RNA-BINDING (RRM) 1.
 FT DOMAIN 184 260 RNA-BINDING (RRM) 2.
 FT DOMAIN 363 437 RNA-BINDING (RRM) 3.
 FT DOMAIN 480 555 RNA-BINDING (RRM) 4.
 FT DOMAIN 342 349 POLY-ALA.
 SQ SEQUENCE 557 AA; 59855 MW; E323F8FBD6DAFF66 CRC64;

Query Match 1.7%; Score 7; DB 1; Length 557;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 284 QFOALLQ 290
 DB 221 QFOALLQ 227

RESULT 62

YHJU_ECOLI STANDARD; PRT; 559 AA.
 ID YHJU_ECOLI
 AC P37859;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein yhjU.
 GN YHJU OR B3538.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=94316500; PubMed=8041620;
 RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
 RT "Analysis of the Escherichia coli genome. V. DNA sequence of the
 region from 76.0 to 81.5 minutes.";
 RL Nucleic Acids Res. 22:2576-2586(1994).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)

DR EMBL; U00039; AAB18516.1; -;
 DR EMBL; AE000431; AAC76563.1; -;
 DR PIR; S47760; S47760.
 DR Ecogene; Egi2265; yhjU.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 559 AA; 62032 MW; 7528D897B0E3E4E9 CRC64;

Query Match 1.7%; Score 7; DB 1; Length 559;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 AATVAAT 141

```
Db 179 AATVAAT 185
|||||
RESULT 63
Y745 ARATH STANDARD; PRT; 597 AA.
AC F59278; Q9C812;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein At1g51745.
GN At1g51745 OR F19C24.5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucots II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizer L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.B., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziani A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uteckebak T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
Thaliana.";
RL Nature 408:816-820 (2000).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT SSP consortium (Salk/Stanford/RGEC).";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 PWMP domain.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to erroneous
CC gene model prediction.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk.)
CC
CC EMBL; AC025294; AAG50888.1; AUT_SEQ.
DR EMBL; BT002989; AAO22798.1; -.
DR PROSITE; PS50812; PWMP; 1.
KW Hypothetical protein.
FT DOMAIN 16 78 PWMP.
FT CONFLICT 268 268 M -> R (IN REF. 1).
SQ SEQUENCE 597 AA; 65044 MW; E59141F5B9DFD79F CRC64;

Query Match 1.7%; Score 7; DB 1; Length 597;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 76 KAKASSS 82

Db 179 AATVAAT 185
|||||
RESULT 64
SYR_PYRHO STANDARD; PRT; 629 AA.
AC O59147;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Arginyl-tRNA synthetase (EC 6.1.1.19) (Arginine-tRNA ligase) (ArgRS).
GN ARG5 OR FHI478.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76 (1998).
CC -1- CATALYTIC ACTIVITY: ATP + L-arginine + tRNA(Arg) = AMP +
CC di-phosphate + L-arginyl-tRNA(Arg).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk.)
CC
CC EMBL; AP000006; BAA30585.1; -.
DR PIR; A71023; A71023.
DR HAMAP; MF 00123; -.
DR InterPro; IPR001278; Arg_tRNA-synt_1c.
DR InterPro; IPR005148; N.
DR Pfam; PF03485; N-Arg; 1.
DR Pfam; PF00750; tRNA-synt 1d; 1.
DR PRINTS; PR01038; TRNASYNTHARG.
DR TIGRFAMs; TIGR00456; argS; 1.
DR PROSITE; PS00178; AA_tRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
FT Complete proteome.
FT SITE 128 138 "HIGH" REGION.
SQ SEQUENCE 629 AA; 72339 MW; 34CE2D841DBAD48A CRC64;

Query Match 1.7%; Score 7; DB 1; Length 629;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 70 LVIMLSK 76
|||||
Db 549 LVIMLSK 555

RESULT 65
COG6_DROME STANDARD; PRT; 630 AA.
ID COG6_DROME
AC Q9V564; Q95RW7;
DT 28-FEB-2003 (Rel. 41, Created)
```

DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Putative conserved oligomeric Golgi complex component 6.
 GN C91968.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J.H., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostain D., Houston K.A., Howland T.J., Wei M.-H., Ithegam C.,
 RA Jaiswal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP REVISIONS, AND ALTERNATIVE SPLICING.
 RC STRAIN=Berkley;
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Battencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley; TISSUE=Embryo;
 RX MEDLINE=22426066; PubMed=12537569;
 RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
 RA George R.A., Guarin H., Krommiller B., Pacleb J.M., Park S., Wan K.H.,
 RA Rubin G.M., Celniker S.E.;
 RT "A Drosophila full-length cDNA resource.";

RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
 CC -!- FUNCTION: Required for normal Golgi function (By similarity).
 CC -!- SUBUNIT: Component of the conserved oligomeric Golgi complex which
 CC is composed of eight different subunits and is required for normal
 CC Golgi morphology and localization (By similarity).
 CC -!- SUBCELLULAR LOCATION: Golgi (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=Q9V564-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=Q9V564-2; Sequence=VSP_001133, VSP_001134;
 CC -!- SIMILARITY: BELONGS TO THE COG6 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AB003834; AAF38956.1; -;
 DR EMBL; AB003834; AAF38956.1; -;
 DR EMBL; AY061082; AAL28630.1; -;
 DR FlyBase; FBgn0033401; CG1968.
 KW Hypothetical protein; Transport; Protein transport; Golgi stack;
 KW Membrane; Alternative splicing.
 FT VARSPLIC 315 333 DISOQLQNALGYADGVCH -> GKIRKCIHYGTVKIVLPVL
 FT VARSPLIC 334 630 Missing (in isoform Short).
 FT VARSPLIC 630 630 Missing (in isoform Short).
 FT VARSPLIC 630 630 Missing (in isoform Short).
 SQ SEQUENCE 630 AA; 71205 MW; 934113C7B266ABFD CRC64;
 Query Match 1.7%; Score 7; DB 1; Length 630;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 333 EGGPGGN 339
 DB 265 EGGPGGN 271
 RESULT 66
 PRIM_UREPA STANDARD; PRT; 641 AA.
 ID AC Q9PPZ6;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA primase (EC 2.7.7.-).
 GN DNAG OR U0494.
 OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.
 OX NCBI_TaxID=134821;
 RN [1]_TaxID=134821;
 RP SEQUENCE FROM N.A.
 RC STRAIN=Serovar 3;
 RX MEDLINE=20500219; PubMed=11048724;
 RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
 RA Cassell G.H.;
 RA "The complete sequence of the mucosal pathogen Ureaplasma
 RT urealyticum.";
 RT Nature 407:757-762(2000).
 RL Nature 407:757-762(2000).
 CC -!- FUNCTION: DNA PRIMASE IS THE POLYMERASE THAT SYNTHESIZES SMALL
 CC RNA PRIMERS FOR THE OKAZAKI FRAGMENTS ON BOTH TEMPLATE STRANDS AT
 CC REPLICATION FORKS DURING CHROMOSOMAL DNA SYNTHESIS.
 CC -!- COPACITOR: BINDS ONE ZINC ION PER MOLECULE (By similarity).
 CC -!- SUBUNIT: Monomer (By similarity).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration -
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AE002147; AAF30906.1; --
 DR HSFP; Q9X4D0; IDQ0.
 DR InterPro; IPR006295; DNA_primase.
 DR InterPro; IPR006171; Toprim_dom.
 DR InterPro; IPR006647; Toprim_dom.
 DR InterPro; IPR006154; Toprim_sub.
 DR InterPro; IPR002894; Znf_CHC2.
 DR Pfam; PF01751; Toprim; 1.
 DR Pfam; PF01807; zf-CHC2; 1.
 DR ProDom; PD002276; Toprim_primase; 1.
 DR ProDom; PD002988; Znf_CHC2; 1.
 DR SMART; SM00493; TOPRIM; 1.
 DR TIGRPFAMs; TIGR01391; dnaG; 1.
 DR Transferase; DNA replication; DNA-directed RNA polymerase; Primosome;
 KW Zinc-finger; Zinc; Metal-binding; Complete proteome.
 FT ZN FING 39 63 CHC2-TYPE (BY SIMILARITY).
 SQ SEQUENCE 641 AA; 75390 MW; A6920168DB58EF47 CRC64;

Query Match 1.7%; Score 7; DB 1; Length 641;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 ENSFLV1 72
 |||||
 DB 191 ENSFLV1 197

RESULT 67
 VP40 HSVEB
 ID VP40 HSVEB STANDARD; PRT; 646 AA.
 AC P28936; Q69263;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Capsid protein P40 (Viron structural gene 35 protein) [Contains:
 DE Capsid protein VP24 (Assemblin) (Protease) (EC 3.4.21.97); Capsid
 DE protein VP22A)].
 OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1).
 GN Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=31520;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=92295566; PubMed=1318506;
 RX Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
 RT "The DNA sequence of equine herpesvirus-1.";
 RL Virology 189:304-316(1992).
 CC -!- FUNCTION: VP22A IS A COMPONENT OF THE CAPSID CORE INVOLVED IN
 CC PROCESSING AND PACKAGING OF PROGENY DNA. VP24 IS A PROTEASE
 CC WHICH CAN PROTEOLYTICALLY CLEAVE ITSELF AND VP22A AT THE C-
 CC TERMINUS.
 CC -!- CATALYTIC ACTIVITY: Cleaves -Ala-|-Ser- and -Ala-|-Ala- bonds in
 CC the scaffold protein.
 CC -!- SIMILARITY: BELONGS TO OTHER HERPESVIRUSES CAPSID PROTEIN VP40.
 CC -!- SIMILARITY: TO OTHER HERPESVIRUSES CAPSID PROTEIN VP40.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; M86664; AAB02470.1; --
 DR EMBL; M86664; AAB02471.1; --

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; I36798; WZBEC8.
 DR HSFP; P16753; IWPO.
 DR MEROPS; S21.001; --
 DR InterPro; IPR001847; Assemblin.
 DR Pfam; PF00716; Peptidase_S21; 1.
 DR PRINTS; PR00236; HSCAPSTDP40.
 KW Coat protein; Hydrolase; Serine protease.
 FT CHAIN 1 646 GENE 35 PROTEIN.
 FT CHAIN 318 646 GENE 35.5 PROTEIN.
 FT CHAIN 1 242 COAT PROTEIN VP24.
 FT CHAIN 243 646 COAT PROTEIN VP22A.
 FT PROPEP 2623 646 C-TERMINAL PEPTIDE.
 FT SITE 242 243 CLEAVAGE (BY THE PROTEASE) (PROBABLE).
 FT SITE 622 623 CLEAVAGE (BY THE PROTEASE) (PROBABLE).
 FT ACT SITE 55 55 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT SITE 123 123 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT SITE 142 142 CHARGE RELAY SYSTEM (BY SIMILARITY).
 SQ SEQUENCE 646 AA; 68579 MW; 1C1C28C848116276 CRC64;

Query Match 1.7%; Score 7; DB 1; Length 646;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 QPQAAPA 135
 |||||
 DB 513 QPQAAPA 519

RESULT 68
 PKNI_COREF
 ID PKNI_COREF STANDARD; PRT; 660 AA.
 AC Q8FUI5;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Probable serine/threonine protein kinase CE0033 (EC 2.7.1.37).
 GN CE0033.
 OS Corynebacterium efficiens.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=152794;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
 RA Kwarabavay Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
 RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
 RA Usuda Y., Sugimoto S.;
 RT "The entire genomic sequence of Corynebacterium efficiens YS-314.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -!- SIMILARITY: Contains 3 PASTA domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; AP005214; BAC16843.1; --
 DR InterPro; IPR005543; PASTA.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF03793; PASTA; 3.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00740; PASTA; 3.
 DR SMART; SM00220; S_TKc; 1.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS00101; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
KW ATP-binding; Repeat; Complete proteome.
FT DOMAIN 9 278 PROTEIN KINASE.
FT DOMAIN 377 443 PASTA 1.
FT DOMAIN 444 512 PASTA 2.
FT DOMAIN 513 577 PASTA 3.
FT NP BIND 15 23 ATP (BY SIMILARITY).
FT BINDING 38 38 ATP (BY SIMILARITY).
FT ACT_SITE 136 136 BY SIMILARITY.
SQ SEQUENCE 660 AA; 69646 MW; D33D797EB02D44B1 CRC64;

Query Match 1.7%; Score 7; DB 1; Length 660;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 PAATAVA 140
Db 318 PAATAVA 324
|||||
318 PAATAVA 324

RESULT 69
SCRT DROME
ID SCRT DROME STANDARD; PRT; 666 AA.
AC P45843;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Scarlet protein.
GN ST.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Canton-S.
RA Garcia R.L., Perkins H.D., Howells A.J.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 200-306 FROM N.A.
RX MEDLINE=89339145; PubMed=2503416;
RA Tearle R.G., Belote J.M., McKeown M., Baker B.S., Howells A.J.;
RT "Cloning and characterization of the scarlet gene of Drosophila melanogaster."
RL Genetics 122:595-606(1989).
CC -1- FUNCTION: PART OF A MEMBRANE-SPANNING PERMEASE SYSTEM NECESSARY FOR THE TRANSPORT OF PIGMENT PRECURSORS INTO PIGMENT CELLS RESPONSIBLE FOR EYE COLOR. SCARLET AND WHITE DIMERIZE FOR THE TRANSPORT OF TRYPTOPHAN.
CC -1- SUBUNIT: HETERODIMER OF SCARLET WITH WHITE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; U39739; AAA82056.1; -
DR EMBL; X76201; CAA53794.1; -
DR FlyBase; FBgn0003515; st.
DR GO; GO:0006727; P:omochrome biosynthesis; IMP.
DR InterPro; IPR003593; AAA_Atpase.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR005284; Pigment_permease.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.

DR SMART; SM00382; AAA; 1.
DR TIGRfams; TIGR00955; 3a01204; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW Pigment; ATP-binding; Transmembrane; Transport.
FT NP BIND 108 115 ATP (POTENTIAL).
FT TRANSMEM 418 438 POTENTIAL.
FT TRANSMEM 445 465 POTENTIAL.
FT TRANSMEM 491 511 POTENTIAL.
FT TRANSMEM 519 539 POTENTIAL.
FT TRANSMEM 552 572 POTENTIAL.
FT TRANSMEM 577 597 POTENTIAL.
FT TRANSMEM 640 660 POTENTIAL.
SQ SEQUENCE 666 AA; 74506 MW; 6796ED4084B59CE4 CRC64;

Query Match 1.7%; Score 7; DB 1; Length 666;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 251 LNLFPQG 257
Db 462 LNLFPQG 468
|||||
462 LNLFPQG 468

RESULT 70
TUP1_KLULA
ID TUP1_KLULA STANDARD; PRT; 682 AA.
AC P56094;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transcriptional repressor TUP1.
GN TUP1.
OS Kluyveromyces lactis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=28985;
RN [1]
RP SEQUENCE FROM N.A.
RA Braun B.R., Johnson A.D.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=20036803; PubMed=10567571;
RA Mukai Y., Matsuo E., Roth S.Y., Harashima S.;
RT "Conservation of histone binding and transcriptional repressor functions in a Schizosaccharomyces pombe Tup1 homolog."
RL Mol. Cell. Biol. 19:8461-8468(1999).
CC -1- FUNCTION: REPRESSSES TRANSCRIPTION BY RNA POLYMERASE II.
CC -1- SIMILARITY: Contains 7 WD repeats.
CC -1- SIMILARITY: BELONGS TO THE TUP1 FAMILY OF WD-REPEAT PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; AF005740; AAB63194.1; -
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 7.
DR ProDom; PD000018; WD40; 4.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS00678; WD_REPEATS_1; 5.
DR PROSITE; PS00082; WD_REPEATS_2; 5.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Transcription regulation; Repressor; Repeat; WD repeat.
FT REPEAT 319 359 WD 1.
FT REPEAT 404 443 WD 2.
FT REPEAT 446 485 WD 3.
FT REPEAT 487 527 WD 4.

FT REPEAT 537 576 WD 5.
 FT REPEAT 586 625 WD 6.
 FT REPEAT 628 673 WD 7.
 FT DOMAIN 172 180 POLY-GLN.
 FT DOMAIN 184 194 POLY-GLN.
 SQ SEQUENCE 682 AA; 74054 MW; E17C2A436232445E CRC64;

Query Match 1.7%; Score 7; DB 1; Length 682;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 ASTATTA 90
 |||||
 Db 233 ASTATTA 239

RESULT 71
 EFG_BRUME
 ID_EFG_BRUME STANDARD; PRT; 694 AA.
 AC Q8YHP3;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Elongation factor G (EF-G).
 GN FUSA OR EME10754.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29459;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / Biotype 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA DelVecchio V.G., Kapatal V., Redkar R.J., Patra G., Muijer C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
 RA Haeckl R., Kyrpides N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 Brucella melitensis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 CC -1- FUNCTION: This protein promotes the GTP-dependent translocation of
 the nascent protein chain from the A-site to the P-site of the
 ribosome.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC EF-G/EF-2 SUBFAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.ebi.ac.uk/ebis/announcement/>
 or send an email to license@ebi.ac.uk).
 CC EMBL: AE009518; AAL51935.1; --
 CC PIR: AD3346; AD3346.
 CC HAWAP: MF_00054; --; 1.
 CC InterPro: IPR004540; EF-G.
 CC InterPro: IPR000795; EF GTPbind.
 CC InterPro: IPR000640; EFG_C.
 CC InterPro: IPR005517; EFG_IV.
 CC InterPro: IPR004161; EFTU_D2.
 CC InterPro: IPR005225; Small GTP.
 CC Pfam: PF00679; EFG_C; 1.
 CC Pfam: PF03764; EFG_IV; 1.
 CC Pfam: PF00009; GTP_EFTU; 1.
 CC Pfam: PF03144; GTP_EFTU_D2; 1.
 CC TIGRFAMs: TIGR00484; EF-G; 1.
 CC TIGRFAMs: TIGR00231; small GTP; 1.
 CC PROSITE: PS00301; EFACITOR GTP; 1.
 KW Elongation factor; Protein biosynthesis; GTP-binding;

KW Complete proteome.
 FT NP_BIND 17 24 GTP (BY SIMILARITY).
 FT NP_BIND 86 90 GTP (BY SIMILARITY).
 FT NP_BIND 140 143 GTP (BY SIMILARITY).
 SQ SEQUENCE 694 AA; 76295 MW; ED1303A991DC34E2 CRC64;

Query Match 1.7%; Score 7; DB 1; Length 694;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 232 AVALPVQ 238
 |||||
 Db 162 AVALPVQ 168

RESULT 72
 STE11_YEAST
 ID_STE11_YEAST STANDARD; PRT; 738 AA.
 AC P23561;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Serine/threonine-protein kinase STE11 (EC 2.7.1.37).
 GN STE11 OR YLR362W OR YLR039.10.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE OF 22-738 FROM N.A.
 RX MEDLINE=91115076; PubMed=2276621;
 RA Rhodes N., Connell L., Errede B.;
 RT "STE11 is a protein kinase required for cell-type-specific
 transcription and signal transduction in yeast.";
 RL Genes Dev. 4:1862-1874(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX MEDLINE=97313267; PubMed=9169871;
 RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Anseorge W.,
 RA Benes V., Brueckner M., Bellus H., Dubois E., Duesterhoft A.,
 RA Entian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
 RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
 RA Louis E.J., Messenguy P., Mewes H.-W., Miesing T., Moestl D.,
 RA Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
 RA Portetelle D., Purnelle B., Reckmann S., Rieger M., Rinke M., Rose M.,
 RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz P.,
 RA Underwood A.P., Urrestarazu L.A., Vandenbol M., Verhaesselt P.,
 RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
 RA Wedler H., Zimmermann P.K., Zollner A., Hani J., Hoheisel J.D.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
 RL Nature 387:87-90(1997).
 RN [3]
 RP POSSIBLE FUNCTION.
 RX MEDLINE=92313935; PubMed=1628833;
 RA Cairns B.R., Ramer S.W., Kornberg K.D.;
 RT "Order of action of components in the yeast pheromone response
 pathway revealed with a dominant allele of the STE11 kinase and the
 multiple phosphorylation of the STE7 kinase.";
 RL Genes Dev. 6:1305-1318(1992).
 CC -1- FUNCTION: SERINE/THREONINE PROTEIN KINASE REQUIRED FOR CELL-TYPE-
 SPECIFIC TRANSCRIPTION AND SIGNAL TRANSDUCTION IN YEAST. IT
 IS THOUGHT THAT IT PHOSPHORYLATES THE STE7 PROTEIN KINASE WHICH
 ITSELF, PHOSPHORYLATES THE FUS3 AND OR KSS1 KINASES.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- PATHWAY: RESPONSE TO PHEROMONE-INDUCED SIGNAL.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE KINASE SUBFAMILY.
 CC -1- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.
 CC -----
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X53431; CAA37522.1; -.
CC DR EMBL; U19103; AAB67571.1; -.
CC DR PIR; S51380; S51380.
CC DR SGD; S0004354; STE11.
CC DR GO; GO:0005737; Cytoplasm; IDA.
CC DR InterPro; IPR000719; Cytosol; IDA.
CC DR InterPro; IPR001660; SAM.
CC DR InterPro; IPR002290; Ser_thr_kinase.
CC DR InterPro; IPR001245; Tyr_kinase.
CC DR Pfam; PF00069; pkinase; 1.
CC DR Pfam; PF00536; SAM; 1.
CC DR PRINTS; PR00109; TYRKINASE.
CC DR ProDom; PD000001; Prot_kinase; 1.
CC DR SMART; SM00220; S_TKc; 1.
CC DR SMART; SM00454; SAM; 1.
CC DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
CC DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
CC DR PROSITE; PS00105; SAM DOMAIN; 1.
CC DR Transferase; Serine/threonine-protein kinase; ATP-binding;
CC KW Phosphonome response.
CC FT DOMAIN 41 105 SAM.
CC FT DOMAIN 436 733 PROTEIN KINASE.
CC FT NP_BIND 442 450 ATP (BY SIMILARITY).
CC FT BINDING 465 465 ATP (BY SIMILARITY).
CC FT ACT_SITE 600 600 BY SIMILARITY.
CC SQ SEQUENCE 738 AA; 83217 MW; A50D69800B346A41 CRC64;

Query Match 1.74; Score 7; DB 1; Length 738;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 86 TATTAKA 92
Db 653 TATTAKA 659
|||||

RESULT 73
ELD_CHICK STANDARD; PRT; 750 AA.
AC P07916;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Elastin precursor (Tropoelastin) (Fragment).
GN ELN.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87242320; PubMed=3593675;
RA Bressan G.M., Argos P., Stanley K.K.;
RT "Repeating structure of chick tropoelastin revealed by complementary
RT DNA cloning."
RL Biochemistry 26:1497-1503(1987).
RN [2]
RP SEQUENCE OF 85-750 FROM N.A. (ISOFORM 2).
RX MEDLINE=88309083; PubMed=2841924;
RA Baule V.J., Foster J.A.;
RT "Multiple chick tropoelastin mRNAs."
RL Biochem. Biophys. Res. Commun. 154:1054-1060(1988).
RN [3]
RP SEQUENCE OF 457-750 FROM N.A.
RC TISSUE=Aorta;
RX MEDLINE=87297534; PubMed=3502711;

```

```

RA Tokimatsu I., Tajima S., Nishikawa T., Tajima M., Fukasawa T.;
RT "Sequence analysis of elastin cDNA from chick aorta and
RT tissue-specific transcription of the elastin gene in developing chick
RT embryo."
RL Arch. Biochem. Biophys. 256:455-461(1987).
CC -!- FUNCTION: MAJOR STRUCTURAL PROTEIN OF TISSUES SUCH AS AORTA AND
CC NUCAL LIGAMENT, WHICH MUST EXPAND RAPIDLY AND RECOVER COMPLETELY.
CC -!- SUBUNIT: THE POLYMERIC ELASTIN CHAINS ARE CROSS-LINKED TOGETHER
CC INTO AN EXTENSIBLE 3D NETWORK.
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX OF ELASTIC FIBERS.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=1;
CC IsoId=P07916-1; Sequence=Displayed;
CC Name=2; Synonyms=Embryonic;
CC IsoId=P07916-2; Sequence=VSP 004241, VSP 004242;
CC -!- PTM: THE CROSSLINKS ARE MADE OF DEAMINATED LYS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M18633; AAA48761.1; -.
CC DR EMBL; M21880; AAA49082.1; -.
CC DR EMBL; M15889; AAA49108.1; -.
CC DR PIR; A26601; A26601.
CC DR InterPro; IPR000087; Collagen.
CC DR InterPro; IPR003979; tropoelastin.
CC Pfam; PF01391; Collagen; 1.
CC PRINTS; PR01500; TROPOELASTIN.
CC KW Structural protein; Connective tissue; Repeat; Signal;
CC Alternative splicing.
CC FT NON_TER 1 1
CC FT SIGNAL <1 24 ELASTIN.
CC FT CHAIN 25 750 8 X TANDEM REPEATS.
CC FT DOMAIN 83 686 1.
CC FT REPEAT 83 127 2.
CC FT REPEAT 219 262 3.
CC FT REPEAT 263 318 4.
CC FT REPEAT 319 393 5.
CC FT REPEAT 394 482 6.
CC FT REPEAT 483 554 7.
CC FT REPEAT 555 619 8.
CC FT REPEAT 620 686 8.
CC FT DISULFID 739 745 BY SIMILARITY.
CC FT MOD_RES 63 63 OXIDATIVE DEAMINATION (POTENTIAL).
CC FT MOD_RES 66 66 OXIDATIVE DEAMINATION (POTENTIAL).
CC FT MOD_RES 111 111 OXIDATIVE DEAMINATION (POTENTIAL).
CC FT MOD_RES 115 115 OXIDATIVE DEAMINATION (POTENTIAL).
CC FT MOD_RES 156 156 OXIDATIVE DEAMINATION (POTENTIAL).
CC FT MOD_RES 159 159 OXIDATIVE DEAMINATION (POTENTIAL).
CC FT MOD_RES 198 198 OXIDATIVE DEAMINATION (POTENTIAL).
CC FT MOD_RES 200 200 OXIDATIVE DEAMINATION (POTENTIAL).
CC FT MOD_RES 235 235 OXIDATIVE DEAMINATION (POTENTIAL).
CC FT MOD_RES 252 252 OXIDATIVE DEAMINATION (POTENTIAL).
CC FT MOD_RES 256 256 OXIDATIVE DEAMINATION (POTENTIAL).
CC FT MOD_RES 297 297 OXIDATIVE DEAMINATION (POTENTIAL).
CC FT MOD_RES 301 301 OXIDATIVE DEAMINATION (POTENTIAL).
CC FT MOD_RES 354 354 OXIDATIVE DEAMINATION (POTENTIAL).
CC FT MOD_RES 357 357 OXIDATIVE DEAMINATION (POTENTIAL).
CC FT MOD_RES 427 427 OXIDATIVE DEAMINATION (POTENTIAL).
CC FT MOD_RES 431 431 OXIDATIVE DEAMINATION (POTENTIAL).
CC FT MOD_RES 513 513 OXIDATIVE DEAMINATION (POTENTIAL).
CC FT MOD_RES 517 517 OXIDATIVE DEAMINATION (POTENTIAL).
CC FT MOD_RES 520 520 OXIDATIVE DEAMINATION (POTENTIAL).
CC FT MOD_RES 586 586 OXIDATIVE DEAMINATION (POTENTIAL).
CC FT MOD_RES 590 590 OXIDATIVE DEAMINATION (POTENTIAL).
CC FT MOD_RES 593 593 OXIDATIVE DEAMINATION (POTENTIAL).

```

FT MOD_RES 655 655 OXIDATIVE DEAMINATION (POTENTIAL).
 FT MOD_RES 658 658 OXIDATIVE DEAMINATION (POTENTIAL).
 FT MOD_RES 719 719 OXIDATIVE DEAMINATION (POTENTIAL).
 FT MOD_RES 722 722 OXIDATIVE DEAMINATION (POTENTIAL).
 FT MOD_RES 743 743 OXIDATIVE DEAMINATION (POTENTIAL).
 FT MOD_RES 748 748 OXIDATIVE DEAMINATION (POTENTIAL).
 FT MOD_RES 750 750 OXIDATIVE DEAMINATION (POTENTIAL).
 FT VARSPLIC 212 212 isoform 2).
 FT G -> GLGFGGQQGVPLGYPIKAPKLP (in isoform 2).
 FT /FTid=VSP 004241.
 FT G -> GVGVGGVGVPG (in isoform 2).
 FT /FTid=VSP 004242.
 FT A -> G (IN REF. 3).
 FT CONFLICT 536 536 G -> A (IN REF. 3).
 FT CONFLICT 571 571 P -> A (IN REF. 3).
 FT CONFLICT 610 610 P -> A (IN REF. 3).
 FT CONFLICT 654 654 A -> R (IN REF. 3).
 FT CONFLICT 667 667 P -> R (IN REF. 3).
 FT CONFLICT 750 750 P -> R (IN REF. 3).
 FT SEQUENCE 750 AA; 63697 MW; E57ECD60CSE556P CRC64;
 SQ
 Query Match 1.7%; Score 7; DB 1; Length 750;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 266 GVVPGAG 272
 DB 279 GVVPGAG 285
 RESULT 74
 ID LYS4_EMENT STANDARD; PRT; 775 AA.
 AC Q92412;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Homoacitase, mitochondrial precursor (EC 4.2.1.36) (Homoacitase hydrolase).
 GN LYSF.
 OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; Emericella.
 OX NCBI_TaxID=164425;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97411901; PubMed=9268014;
 RA Weidner G., Steffan B., Brakhage A.A.;
 RT "The Aspergillus nidulans lysF gene encodes homoacitase, an enzyme involved in the fungus-specific lysine biosynthesis pathway.";
 RL Mol. Gen. Genet. 255:237-247(1997).
 CC -1- FUNCTION: RESPONSIBLE FOR THE DEHYDRATION OF CIS-HOMOACITATE TO HOMOISOCITRIC ACID.
 CC -1- CATALYTIC ACTIVITY: 2-hydroxybutane-1,2,4-tricarboxylate = but-1-ene-1,2,4-tricarboxylate + H(2)O.
 CC -1- COFACTOR: Binds 1 4Fe-4S cluster per subunit (By similarity).
 CC -1- PATHWAY: Lysine biosynthesis; alpha-aminoacidic acid pathway; third step.
 CC -1- SIMILARITY: BELONGS TO THE ACONITASE/IPM ISOMERASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; X99624; CAA67943.1; -
 CC InterPro; IPR000573; Aconitase C.
 CC InterPro; IPR001030; Aconitase N.
 CC InterPro; IPR004418; Homoacitase.
 CC Pfam; PF00330; aconitase; 1.
 CC Pfam; PF00694; Aconitase C; 1.
 CC PRINTS; PR00415; ACONITASE.

DR ProDom; PD000511; Aconitase N; 1.
 DR TIGRfam; TIGR00139; h_aconitase; 1.
 DR PROSITE; PS00450; ACONITASE 1; 1.
 DR PROSITE; PS01244; ACONITASE 2; 1.
 KW Lysine biosynthesis; Lysine; Mitochondrion; Transit peptide;
 KW Iron-sulfur.
 FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
 FT CHAIN ? 775 HOMOACONITASE.
 FT METAL 394 394 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 FT METAL 462 462 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 FT METAL 465 465 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 FT SEQUENCE 775 AA; 84038 MW; ABC08D6B01169632 CRC64;
 SQ
 Query Match 1.7%; Score 7; DB 1; Length 775;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 347 AVPQTLT 353
 DB 63 AVPQTLT 69
 RESULT 75
 ID FRZE_MYXXA STANDARD; PRT; 777 AA.
 AC P18769;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Gliding motility regulatory protein (EC 2.7.3.-).
 GN FRZE.
 OS Myxococcus xanthus.
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
 OC Cytophactereae; Myxococcaceae; Myxococcus.
 OX NCBI_TaxID=34;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90332690; PubMed=2165608;
 RA McCleary W.R., Zusman D.R.;
 RT "Frze of Myxococcus xanthus is homologous to both Chea and Chey of Salmoneella typhimurium.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:5898-5902(1990).
 RN [2]
 RP PHOSPHORYLATION OF HIS-49.
 RX MEDLINE=91072208; PubMed=2123853;
 RA McCleary W.R., Zusman D.R.;
 RT "Purification and characterization of the Myxococcus xanthus Frze protein shows that it has autophosphorylation activity.";
 RL J. Bacteriol. 172:6661-6668(1990).
 CC -1- FUNCTION: FRZE IS INVOLVED IN A SENSORY TRANSDUCTION PATHWAY THAT CONTROLS THE FREQUENCY AT WHICH CELLS REVERSE THEIR GLIDING DIRECTION. FRZE SEEMS TO BE CAPABLE OF AUTOPHOSPHORYLATING ITSELF ON AN HISTIDINE RESIDUE AND THEN TO TRANSFER THAT GROUP TO AN ASPARTATE RESIDUE IN THE C-TERMINAL PART OF THE PROTEIN.
 CC -1- SIMILARITY: Contains 1 cheW-like domain.
 CC -1- SIMILARITY: Contains 1 histidine kinase domain.
 CC -1- SIMILARITY: Contains 1 HPT domain.
 CC -1- SIMILARITY: Contains 1 response regulatory domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; M35192; AAA25396.1; -
 CC InterPro; IPR000573; Aconitase C.
 CC InterPro; IPR001030; Aconitase N.
 CC InterPro; IPR004418; Homoacitase.
 CC Pfam; PF00330; aconitase; 1.
 CC Pfam; PF00694; Aconitase C; 1.
 CC PRINTS; PR00415; ACONITASE.


```

DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR002570; Hpt.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF01584; Chew; 1.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF01827; Hpt; 1.
DR Pfam; PF00072; response_reg; 1.
DR PRINTS; PR00344; ECTRLSENSOR.
DR ProDom; PD000039; Response_reg; 1.
DR SMART; SM00260; Chew; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00073; HPT; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS00851; CHEW; 1.
DR PROSITE; PS0109; HIS_KIN; 1.
DR PROSITE; PS50894; HPT; 1.
DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
KW Sensory transduction; Transferrase; Kinase; Phosphorylation.
FT DOMAIN 1 108 HPT.
FT DOMAIN 270 509 HISTIDINE KINASE.
FT DOMAIN 511 645 CHEW-LIKE.
FT DOMAIN 660 776 RESPONSE REGULATORY.
FT MOD_RES 49 49 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT DOMAIN 130 197 ALA/PRO-RICH (POSSIBLE HINGE REGION).
SQ SEQUENCE 777 AA; 83189 MW; 9912BD40991C69E5 CRC64;

Query Match 1.7%; Score 7; DB 1; Length 777;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 PTOAPVA 119
Db 175 PTOAPVA 181

```

Search completed: December 17, 2003, 06:33:48
 Job time : 15.5265 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 06:26:06 ; Search time 39.8189 Seconds
(without alignments)
2624.667 Million cell updates/sec

Title: US-09-805-550-2

Perfect score: 405

Sequence: 1 MKNVTKLTGTFEIEASPD.....BELTANYLLDHGHEFDQQQ 405

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

SPTREMBL_23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	5.2	378	10 Q94C35	Q94C35 arabidopsis
2	21	5.2	378	10 Q9FF16	Q9FF16 arabidopsis
3	20	4.9	389	10 Q9STA6	Q9STA6 lycopersico
4	18	4.4	113	10 Q9SA20	Q9SA20 arabidopsis
5	18	4.4	365	10 Q94CE9	Q94CE9 arabidopsis
6	18	4.4	371	10 Q8LA46	Q8LA46 arabidopsis
7	17	4.2	382	10 Q03990	Q03990 daucus caro
8	14	3.5	392	10 Q04742	Q04742 oryza sativ
9	13	3.2	246	10 Q9S9L8	Q9S9L8 arabidopsis
10	13	3.2	367	10 Q9MA10	Q9MA10 arabidopsis
11	13	3.2	379	10 Q03991	Q03991 daucus caro
12	13	3.2	419	10 Q9M887	Q9M887 arabidopsis
13	10	2.5	65	10 Q9SCA8	Q9SCA8 lycopersico
14	10	2.5	1860	5 Q9UI29	Q9UI29 leishmania
15	9	2.2	269	5 Q9VI97	Q9VI97 drosophila
16	9	2.2	286	16 Q8P454	Q8P454 xanthomonas

17	9	2.2	365	10 Q9AWS0	Q9AWS0 oryza sativ
18	9	2.2	891	11 Q8VHG2	Q8VHG2 mus musculus
19	9	2.2	3069	16 P95029	P95029 mycobacteri
20	8	2.0	72	2 Q48506	Q48506 lactococcus
21	8	2.0	192	16 Q8DB85	Q8DB85 vibrio vuln
22	8	2.0	227	7 Q9TP09	Q9TP09 brachydanio
23	8	2.0	248	16 Q8XRR4	Q8XRR4 ralstonia s
24	8	2.0	254	16 Q8ZB70	Q8ZB70 yersinia pe
25	8	2.0	255	16 Q8DC23	Q8DC23 vibrio vuln
26	8	2.0	256	16 Q9KPF3	Q9KPF3 vibrio chol
27	8	2.0	264	2 Q52007	Q52007 vibrio fisc
28	8	2.0	272	10 Q80422	Q80422 oryza sativ
29	8	2.0	273	10 Q40628	Q40628 oryza sativ
30	8	2.0	278	16 Q8DIB5	Q8DIB5 yersinia pe
31	8	2.0	289	2 Q9X2P4	Q9X2P4 mycobacteri
32	8	2.0	300	10 Q65002	Q65002 betula verr
33	8	2.0	308	10 Q9FUM6	Q9FUM6 betula verr
34	8	2.0	322	7 Q8HWF2	Q8HWF2 brachydanio
35	8	2.0	339	7 Q9GJK0	Q9GJK0 salmo trutt
36	8	2.0	354	5 P91201	P91201 caenorhabdi
37	8	2.0	355	16 Q8PNM4	Q8PNM4 xanthomonas
38	8	2.0	356	16 Q8Y277	Q8Y277 ralstonia s
39	8	2.0	396	9 Q9XJA9	Q9XJA9 streptococc
40	8	2.0	415	10 Q9FNS5	Q9FNS5 chlamydomon
41	8	2.0	425	11 Q9DBZ2	Q9DBZ2 mus musculu
42	8	2.0	462	17 Q9HP89	Q9HP89 halobacteri
43	8	2.0	466	16 Q9RK92	Q9RK92 streptomyce
44	8	2.0	467	10 Q8RVT6	Q8RVT6 pisum sativ
45	8	2.0	473	16 Q92KE5	Q92KE5 rhizobium m
46	8	2.0	480	10 Q8S0B0	Q8S0B0 oryza sativ
47	8	2.0	514	16 Q8Y1V5	Q8Y1V5 ralstonia s
48	8	2.0	566	16 P72770	P72770 synecocyst
49	8	2.0	571	5 Q9VUD3	Q9VUD3 drosophila
50	8	2.0	595	2 Q85688	Q85688 caulobacter
51	8	2.0	643	16 P72058	P72058 mycobacteri
52	8	2.0	653	16 Q8YGD6	Q8YGD6 brucella me
53	8	2.0	687	16 Q9ABQ3	Q9ABQ3 caulobacter
54	8	2.0	695	11 Q8CI48	Q8CI48 mus musculu
55	8	2.0	704	16 Q8G116	Q8G116 brucella su
56	8	2.0	757	16 Q8VIT6	Q8VIT6 mycobacteri
57	8	2.0	898	11 Q8KLS4	Q8KLS4 mus musculu
58	8	2.0	898	11 Q08721	Q08721 rattus norv
59	8	2.0	908	16 Q8CJX8	Q8CJX8 streptomyce
60	8	2.0	944	11 Q922S3	Q922S3 mus musculu
61	8	2.0	1327	2 Q9X7M2	Q9X7M2 staphylococ
62	8	2.0	1665	11 Q8VII1	Q8VII1 mus musculu
63	8	2.0	1795	16 Q9LCJ9	Q9LCJ9 staphylococ
64	8	2.0	2478	2 Q9RL69	Q9RL69 staphylococ
65	8	2.0	2478	2 Q9LCH2	Q9LCH2 staphylococ
66	8	2.0	2481	16 Q99QR6	Q99QR6 staphylococ
67	8	2.0	2593	12 Q8QL53	Q8QL53 sleeping di
68	8	2.0	2861	5 Q9UIC3	Q9UIC3 leishmania
69	7	1.7	48	6 Q95M44	Q95M44 bos taurus
70	7	1.7	58	16 Q9K8K3	Q9K8K3 bacillus ha
71	7	1.7	74	7 P79662	P79662 oncorhynch
72	7	1.7	74	7 P79652	P79652 oncorhynch
73	7	1.7	74	16 Q8XTY0	Q8XTY0 ralstonia s
74	7	1.7	75	10 Q8L656	Q8L656 oryza sativ
75	7	1.7	83	2 Q9L824	Q9L824 enterococc
76	7	1.7	85	2 Q54454	Q54454 streptococc
77	7	1.7	85	10 Q9XB17	Q9XB17 alexandrium
78	7	1.7	87	16 Q8YV58	Q8YV58 anabaena sp
79	7	1.7	89	15 Q85559	Q85559 baboon endo
80	7	1.7	96	12 Q9QBR1	Q9QBR1 myxoma viru
81	7	1.7	96	12 Q9Q336	Q9Q336 shope fibro
82	7	1.7	100	16 Q99SY0	Q99SY0 staphylococ
83	7	1.7	100	16 Q8CRT7	Q8CRT7 staphylococ
84	7	1.7	105	2 Q8KYV0	Q8KYV0 uncultured
85	7	1.7	107	10 Q8GUE6	Q8GUE6 cyanophora
86	7	1.7	108	16 Q8Z6L4	Q8Z6L4 salmonella
87	7	1.7	108	16 Q84944	Q84944 salmonella
88	7	1.7	109	10 Q8H8Y5	Q8H8Y5 oryza sativ
89	7	1.7	113	2 Q933S0	Q933S0 campylobact

90 7 1.7 113 11 Q9D5N3
 91 7 1.7 114 10 Q9ATF7
 92 7 1.7 115 10 Q9ATF8
 93 7 1.7 118 10 Q9ATF9
 94 7 1.7 119 10 Q9SJP7
 95 7 1.7 119 10 Q8GY07
 96 7 1.7 126 2 P72563
 97 7 1.7 128 16 Q8X309
 98 7 1.7 129 16 Q9XA01
 99 7 1.7 134 9 Q8SCX6
 100 7 1.7 134 16 Q98NR2

ALIGNMENTS

RESULT 1

Q94C35 PRELIMINARY; PRT; 378 AA.
 ID Q94C35
 AC Q94C35;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical 40.1 kDa protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC euroids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Shinn P.,
 RA Tracy S.E., Banh J., Bowser L., Carninci P., Chung M.K., Kamiya A.,
 RA Goldsmith A.D., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,
 RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
 RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Pham P.K., Quach H.L.,
 RA Sakano H., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C.,
 RA Toriumi M., Yamada K., Yu G., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cDNA clones.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
 RT "Full-length messenger RNA sequences greatly improve genome
 RT annotation.";
 RL Genome Biol. 0:0-0 (2002).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RT "Full-length cDNA from Arabidopsis thaliana.";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY037181; AAK59766.1; -;
 DR InterPro; IPR004806; Rad23.
 DR InterPro; IPR006636; ST1.
 DR InterPro; IPR000449; UBA_domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 2.
 DR Pfam; PF00240; ubiquitin; 1.
 DR SMART; SM00727; ST1; 1.
 DR SMART; SM00165; UBA; 2.
 DR SMART; SM00213; UBO; 1.
 DR TIGRfam; TIGR00601; rad23; 1.
 DR PROSITE; PS50053; UBIQUITIN_2; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 378 AA; 40075 MW; E49B591527A3EC27 CRC64;

Query Match 5.2%; Score 21; DB 10; Length 378;
 Best Local Similarity 100.0%; Pred. No. 8.8e-12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

292 VQANPQLQPMQLGKQNPQ 312
 |||||
 269 VQANPQLQPMQLGKQNPQ 289

RESULT 2

Q9FF16 PRELIMINARY; PRT; 378 AA.
 ID Q9FF16
 AC Q9FF16;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE DNA repair protein RAD23 homolog (Hypothetical protein)

DE (At5g38470/At5g38470).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC euroids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RX MEDLINE=97471969; PubMed=9330910;
 RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
 RA Miyajima N., Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
 RT features of the 1.6 Mb regions covered by twenty physically assigned
 RT P1 clones.";
 RL DNA Res. 4:215-230 (1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cDNA clones.";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Shinn P., Banh J.,
 RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
 RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis ORF clones.";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
 RT "Full-length messenger RNA sequences greatly improve genome
 RT annotation.";
 RL Genome Biol. 0:0-0 (2002).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RT "Full-length cDNA from Arabidopsis thaliana.";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB005248; BAB09359.1; -;
 DR EMBL; AY058196; AAL25609.1; -;
 DR EMBL; AY081835; AAL87405.1; -;
 DR EMBL; AY087564; AAM65106.1; -;
 DR HSP; P54725; IDV0.
 DR InterPro; IPR004806; Rad23.
 DR InterPro; IPR000449; UBA domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 2.
 DR Pfam; PF00240; ubiquitin; 1.
 DR TIGRfam; TIGR00601; rad23; 1.
 DR PROSITE; PS50053; UBIQUITIN_2; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 378 AA; 40066 MW; DD9B59153EA3B6A9 CRC64;

Query Match 5.2%; Score 21; DB 10; Length 378;
 Best Local Similarity 100.0%; Pred. No. 8.8e-12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

292 VQANPQLQPMQLGKQNPQ 312

```

Db      269 VOANPQILQPMQLQELGKQNPQ 289
|||||
RESULT 3
Q9STA6 PRELIMINARY; PRT; 389 AA.
AC Q9STA6;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE RAD23 protein.
GN RAD23.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. West Virginia 106; TISSUE=Fruit;
RA Lemaire-Chamley M., Petit J., Raymond P., Chevalier C.;
RT "Analysis of gene expression during early tomato fruit development by
RT mRNA differential display.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ243875; CAB51544.1; -.
DR HSSP; P54725; 1DV0.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006636; STI1.
DR InterPro; IPR000449; UBA_domain.
DR InterPro; IPR000626; UBIQUITIN.
DR Pfam; PF00627; UBA; 2.
DR SMART; SM00727; STI1; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBQ; 1.
DR TIGRPFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS50053; UBIQUITIN 2; 1.
SQ SEQUENCE 389 AA; 41508 MW; D2BAEDF0FE70778A CRC64;

Query Match 4.9%; Score 20; DB 10; Length 389;
Best Local Similarity 100.0%; Pred. No. 8.7e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      292 VOANPQILQPMQLQELGKQNP 311
|||||
Db      281 VOANPQILQPMQLQELGKQNP 300
|||||

RESULT 4
Q9SA20 PRELIMINARY; PRT; 113 AA.
AC Q9SA20;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE F309.1 protein.
GN F309.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Vysotskaia V.S., Schwartz J., Yu G., Toriumi M., Lenz C., Liu S.,
RA Lee J., Li J., Kremenetskaia I., Liu A., Luros J., Gonzalez A.,
RA Altafi H., Araujo R., Chao Q., Conn L., Conway A.B., Dunn P.,
RA Hansen N., Huizar L., Kim C., Palm C., Rowley D., Shinn P., Walker M.,
RA Davis R.W., Ecker J.R., Federpriel N.A., Theologis A.;
RT "The sequence of BAC F309 from Arabidopsis thaliana chromosome 1.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

```

```

RN      [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Theologis A.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006341; AAD34676.1; -.
DR HSSP; P54725; 1DV0.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF00627; UBA; 1.
DR SMART; SM00165; UBA; 1.
SQ SEQUENCE 113 AA; 12947 MW; 8BE7CB15C9E12380 CRC64;

Query Match 4.4%; Score 18; DB 10; Length 113;
Best Local Similarity 100.0%; Pred. No. 2.5e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      295 NPQILQPMQLQELGKQNPQ 312
|||||
Db      5 NPQILQPMQLQELGKQNPQ 22
|||||

RESULT 5
Q94CE9 PRELIMINARY; PRT; 365 AA.
AC Q94CE9;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Putative RAD23 protein.
GN F20B17.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,
RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E.,
RA Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,
RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
RA Shinzaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "Full length cDNA of gene F20B17.8 (GI:7715605).";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinn P., Southwick A., Shinzaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Full length cDNA of gene F20B17.8 (GI:7715605).";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY034912; AAK59419.1; -.
DR EMBL; AY063103; AAL34277.1; -.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006636; STI1.
DR InterPro; IPR000449; UBA_domain.
DR InterPro; IPR000626; UBIQUITIN.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00727; STI1; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBQ; 1.
DR TIGRPFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS50053; UBIQUITIN 2; 1.
SQ SEQUENCE 365 AA; 39157 MW; 4FED9EC59B467745 CRC64;

```

Query Match 4.4%; Score 18; DB 10; Length 365;
 Best Local Similarity 100.0%; Pred. No. 7.7e-09;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 295 NPQILQPMQLQELGKQNPQ 312
 Db 257 NPQILQPMQLQELGKQNPQ 274

RESULT 6

Q8LA46 PRELIMINARY; PRT; 371 AA.
 AC Q8LA46;
 DT 01-OCT-2002 (TREMELrel. 22, Created)
 DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE DNA repair protein RAD23, putative.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eudicots II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
 RA Feldmann K.,
 RA "Full-length cDNA from Arabidopsis thaliana";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY088037; AAM65583.1; --
 DR InterPro; IPR004806; Rad23.
 DR InterPro; IPR006636; STIL.
 DR InterPro; IPR00449; UBA_domain.
 DR Pfam; PF00627; UBA; 2.
 DR SMART; SM00727; STIL; 1.
 DR SMART; SM00165; UBA; 2.
 DR SMART; SM00213; UBQ; 1.
 DR TIGRFAMs; TIGR00601; rad23; 1.
 DR PROSITE; PS50053; UBIQUITIN 2; 1.
 SQ SEQUENCE 371 AA; 39747 MW; 081493086EA976E7 CRC64;

Query Match 4.4%; Score 18; DB 10; Length 371;
 Best Local Similarity 100.0%; Pred. No. 7.8e-09;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 295 NPQILQPMQLQELGKQNPQ 312
 Db 263 NPQILQPMQLQELGKQNPQ 280

RESULT 7

O03990 PRELIMINARY; PRT; 382 AA.
 AC O03990;
 DT 01-JUL-1997 (TREMELrel. 04, Created)
 DT 01-JUL-1997 (TREMELrel. 04, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE RAD23, isoform I.
 OS Daucus carota (Carrot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; campanulids; Apiales; Apiaceae; Daucus.
 OX NCBI_TaxID=4039;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=W001C;

RX MEDLINE=98345997; PubMed=9681019;
 RA Sturm A., Leinhardt S.;
 RT "Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in yeast.";
 RL Plant J. 13:815-821(1998).
 DR EMBL; Y12013; CAA72741.1; --
 DR HSSP; P54725; IDV0.
 DR InterPro; IPR002965; P-rich_extensn.
 DR InterPro; IPR004806; Rad23.
 DR InterPro; IPR006636; STIL.
 DR InterPro; IPR000449; UBA_domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 2.
 DR Pfam; PF00240; ubiquitin; 1.
 DR PRINTS; PR01217; PRICHEXTENS.
 DR SMART; SM00727; STIL; 1.
 DR SMART; SM00165; UBA; 2.
 DR SMART; SM00213; UBQ; 1.
 DR TIGRFAMs; TIGR00601; rad23; 1.
 DR PROSITE; PS50053; UBIQUITIN 2; 1.
 SQ SEQUENCE 382 AA; 40344 MW; 27FF02A2402F3DC2 CRC64;

Query Match 4.2%; Score 17; DB 10; Length 382;
 Best Local Similarity 100.0%; Pred. No. 7.7e-08;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 295 NPQILQPMQLQELGKQNP 311
 Db 282 NPQILQPMQLQELGKQNP 298

RESULT 8

Q40742 PRELIMINARY; PRT; 392 AA.
 AC Q40742;
 DT 01-NOV-1996 (TREMELrel. 01, Created)
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE OSRAD23.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nipponbare;
 RX MEDLINE=97369378; PubMed=9225866;
 RA Schultz T.F., Quatrano R.S.;
 RT "Characterization and expression of a rice RAD23 gene.";
 RL Plant Mol. Biol. 34:557-562(1997).
 DR EMBL; U63530; AAB65841.1; --
 DR HSSP; P54725; IDV0.
 DR Gramene; Q40742; --
 DR InterPro; IPR004806; Rad23.
 DR InterPro; IPR006636; STIL.
 DR InterPro; IPR000449; UBA_domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 2.
 DR Pfam; PF00240; ubiquitin; 1.
 DR SMART; SM00727; STIL; 1.
 DR SMART; SM00165; UBA; 2.
 DR SMART; SM00213; UBQ; 1.
 DR TIGRFAMs; TIGR00601; rad23; 1.
 DR PROSITE; PS50053; UBIQUITIN 2; 1.
 SQ SEQUENCE 392 AA; 41754 MW; BD6E08574CC7CACB CRC64;

Query Match 3.5%; Score 14; DB 10; Length 392;
 Best Local Similarity 100.0%; Pred. No. 7.1e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 VADVKKRIETTTQGO 36
 |||||

```

Db      23 VADVKRIIETQGG 36

RESULT 9
Q9S9L8 PRELIMINARY; PRT; 246 AA.
AC Q9S9L8
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE T24D18.27 protein.
GN T24D18.27.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Liu S.X., Yu G., Sakano H., Jhaveri A., Lee J.M., Lenz C., Pham P.,
RA Toriumi M., Chin C., Chiou J., Choi E., Chung M., Gonzalez A.,
RA Hwang B., Koo T., Li J., Liu A., Vaysberg M., Altafi H., Brooks S.,
RA Buehler E., Chao Q., Conn L., Conway A., Hansen N., Johnson-Hopson C.,
RA Khan S., Kim C., Lam B., Nguyen M., Palm C., Shinn P., Tambunga G.,
RA Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;
RA "The sequence of BAC T24D18 from Arabidopsis thaliana chromosome 1.";
RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Theologis A.;
RA Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC010924; AAF18513.1; -.
DR HSSP; P02248; IUBI.
DR InterPro; IPR000449; UBA domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 1.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00165; UBA; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 246 AA; 25966 MW; 0867D1E7D2B8FFE CRC64;

Query Match 3.2%; Score 13; DB 10; Length 246;
Best Local Similarity 100.0%; Pred. No. 0.00044;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 182 RALRAAYNNPERA 194
Db 155 RALRAAYNNPERA 167
|||||
DE F20B17.8.

RESULT 10
Q9WA10 PRELIMINARY; PRT; 367 AA.
ID Q9WA10
AC Q9WA10
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE F20B17.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Khan S., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Kim C.,
RA Shinn P., Altafi H., Bei Q., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Hwang B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,

```

```

RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F20B17 from chromosome
RT I.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Hwang B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu S., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC010793; AAF68123.1; -.
DR HSSP; P54725; 1DV0.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR000449; UBA domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBQ; 1.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 367 AA; 39790 MW; C6D01134B5C69CB6 CRC64;

Query Match 3.2%; Score 13; DB 10; Length 367;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 182 RALRAAYNNPERA 194
Db 168 RALRAAYNNPERA 180
|||||
DE F20B17.8.

RESULT 11
O03991 PRELIMINARY; PRT; 379 AA.
ID O03991
AC O03991;
DT 01-JUL-1997 (TRENBLrel. 04, Created)
DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE RAD23 protein, isoform II.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Apiales; Apiaceae; Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W001C;
RA MEDLINE=98345997; PubMed=9681019;
RA Sturm A., Leinhard S.;
RT "Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in
RT yeast.";
RL Plant J. 13:815-821(1998).
DR EMBL; Y12014; CAA72742.1; -.
DR HSSP; P54725; 1DV0.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006636; STIL.
DR InterPro; IPR000449; UBA domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.

```

DR Pfam; PF00240; ubiqtutin; 1.
DR SMART; SM00727; STI1; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBAQ; 1.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS0053; UBIQUITIN 2; 1.
SQ SEQUENCE 379 AA; 40530 MW; B26697B439CC5929 CRC64;

Query Match 3.2%; Score 13; DB 10; Length 379;
Best Local Similarity 100.0%; Pred. No. 0.00067;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 182 RALRAAYNNPERA 194
DB 172 RALRAAYNNPERA 184
|||||

RESULT 12

Q9M887 ID Q9M887 PRELIMINARY; PRT; 419 AA.
AC Q9M887;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative RAD23 (AT3g02540/F16B3_17).
GN F16B3.17.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cnv. Columbia;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RA "Arabidopsis thaliana chromosome III BAC F16B3 genomic sequence";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Gibson H.A., Goldsmith A.D.,
RA Hayaishizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yu G., Yu S., Shinozaki K.,
RA Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Shinn P., Banh J., Bowser L.,
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayaishizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yu G., Yu S., Shinozaki K.,
RA Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis ORF clones";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP EMBL; AC021640; AAF32461.1;
DR EMBL; AV039562; AAK62617.1;
DR EMBL; AV113034; AAM47342.1;
DR HSSP; P54725; 1DV0.
DR InterPro; IPR002965; P rich_extensn.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF00627; UBA; 2.
DR PRINTS; PR01217; PRICEXTENSN.

DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS0053; UBIQUITIN 2; 1.
SQ SEQUENCE 419 AA; 44247 MW; CAA13BC4FFEB1E25 CRC64;
Query Match 3.2%; Score 13; DB 10; Length 419;
Best Local Similarity 100.0%; Pred. No. 0.00073;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 TIQQLDMGGGTW 175
DB 190 TIQQLDMGGGTW 202
|||||

RESULT 13

Q9SCA8 ID Q9SCA8 PRELIMINARY; PRT; 65 AA.
AC Q9SCA8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Rad23 protein (Fragment).
GN RAD23.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cnv. West Virginia 106; TISSUE=fruit;
RA Lemaire-Chamley M., Petit J., Causse M., Raymond P., Chevalier C.;
RT "Isolation and characterization of cDNAs expressed during early
development of tomato fruit by mRNA differential display";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ270958; CAB65692.1;
DR HSSP; P54725; 1DV0.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF00627; UBA; 1.
DR SMART; SM00165; UBA; 1.
FT NON_TER 1
SQ SEQUENCE 65 AA; 7233 MW; 0538945EE9030B4D CRC64;

Query Match 2.5%; Score 10; DB 10; Length 65;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 353 TVTPEERBAI 362
DB 16 TVTPEERBAI 25
|||||

RESULT 14

Q9U129 ID Q9U129 PRELIMINARY; PRT; 1860 AA.
AC Q9U129;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Probable kinase.
GN L4738.03.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RA Hilbert H., Wedler H., Duesterhoef A., Ivens A.C.,
RA Murphy L., Quail M., Rajandream M.A., Barrell B.G.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RC STRAIN=98146435; PubMed=9477341;
RX MEDLINE=98146435; PubMed=9477341;

RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RA "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AL135930; CAB64574.1; -!
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1860 AA; 194940 MW; 8843E84349DD7AD CRC64;

Query Match 2.5%; Score 10; DB 5; Length 1860;
Best Local Similarity 100.0%; Pred.No.2.7; 0; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

Qy 99 PAAPVAPAA 108
Db 1596 PAAPVAPAA 1605
|||||

RESULT 15
Q9VI97
ID Q9VI97 PRELIMINARY; PRT; 269 AA.
AC Q9VI97
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG31496 protein.
GN NXF4 OR CG14604 OR CG31501.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Dou L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fowler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.I., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levisky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J.J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Dou L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwan C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of *Drosophila melanogaster* genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Miera S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradscky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of *Drosophila melanogaster* genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RP Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003672; AAF54028.2; -!
DR FlyBase; FBGN0051501; nxf4.
SQ SEQUENCE 269 AA; 27072 MW; D69F5D742788F331 CRC64;

Query Match 2.2%; Score 9; DB 5; Length 269;
Best Local Similarity 100.0%; Pred.No.4.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 99 PAAPVAPAA 107
Db 158 PAAPVAPAA 166
|||||

RESULT 16
Q8P454
ID Q8P454 PRELIMINARY; PRT; 286 AA.
AC Q8P454
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein XCC3862.
GN XCC3862.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;

RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Camarotte G., Canavan F., Cardoso J., Chambergo F., Ciapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Subthal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities";
 RL Nature 417:459-463 (2002).
 DR EMBL; AF012507; RAM43093.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 286 AA; 28621 MW; C884041CE7BE6F54 CRC64;

Query Match 2.2%; Score 9; DB 16; Length 286;
 Best Local Similarity 100.0%; Pred. No. 4.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 PVAPAASVA 110
 DB 135 PVAPAASVA 143

RESULT 17

Q9AWS0 PRELIMINARY; PRT; 365 AA.
 ID Q9AWS0
 AC Q9AWS0
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Putative DNA binding protein RAV2.
 GN P0480E02.24.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]_TaxID=4530;
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 RL clone: P0480E02";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF002913; BAB21218.1; -;
 DR HSSP; O80337; 2GCC.
 DR Gramene; Q9AWS0; -;
 DR InterPro; IPR003340; TF_B3.
 DR InterPro; IPR001471; TF_ERF.
 DR Pfam; PF00847; AP2-domain; 1.
 DR Pfam; PF02362; B3; 1.
 DR PRINTS; PR00367; ETRSPLEMMNT.
 DR ProDom; PD001423; TF_ERF; 1.
 DR SMART; SM00380; AP2; 1.
 SQ SEQUENCE 365 AA; 39280 MW; 0A1EAD4E7986E6C CRC64;

Query Match 2.2%; Score 9; DB 10; Length 365;
 Best Local Similarity 100.0%; Pred. No. 5.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTATT 89
 DB 10 SSGASTATT 18

RESULT 18

Q8VHG2 PRELIMINARY; PRT; 891 AA.
 ID Q8VHG2
 AC Q8VHG2
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Angiotensin.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B6CBAF2; TISSUE=Placenta;
 RA Troyanovsky B., Bratt A., Holmgren L.;
 RT "Mouse angiotensin";
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF461135; AAL73436.1; -;
 DR InterPro; IPR000104; Antifreeze_1.
 DR PRINTS; PR00308; ANTIFREEZE1.
 SQ SEQUENCE 891 AA; 95111 MW; C054452A4CD6088 CRC64;

Query Match 2.2%; Score 9; DB 11; Length 891;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 PAAPVAPAA 107
 DB 713 PAAPVAPAA 721

RESULT 19

P95029 PRELIMINARY; PRT; 3069 AA.
 ID P95029
 AC P95029
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE RV2524C (Fatty-acid synthase).
 GN FAS OR RV2524C OR MT2600 OR MTCY159.32 OR MTV009.09C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]_TaxID=1773;
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badoock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sultston J.B., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z83863; CAB06201.1; -;
 DR EMBL; AE007096; AAK46907.1; -;

```

DR TIGR; MT2600; -.
DR Tuberculin; RV2524C; -.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR003965; Fatty acid synth.
DR InterPro; IPR000794; Ketoacyl-synt.
DR InterPro; IPR002539; MacC dehydratase.
DR Pfam; PF00698; Acyl_transf; 1.
DR Pfam; PF00109; ketoacyl-synt; 1.
DR Pfam; PF02801; ketoacyl-synt_C; 1.
DR Pfam; PF01575; MacC dehydratase; 1.
DR PRINTS; PR01483; FASYNTHASE.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
KW COMPLETE proteome.
FT CONFLICT 2699 2699 A -> S (IN REF. 2).
SQ SEQUENCE 3069 AA; 326251 MW; 25EF022D54943D4C CRC64;

Query Match 2.2%; Score 9; DB 16; Length 3069;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 AAPVAPAAAS 108
Db 1754 AAPVAPAAAS 1762
|||||

RESULT 20
Q48506 PRELIMINARY; PRT; 72 AA.
AC Q48506;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Beta-lactamase (Fragment).
GN BETA-LACTAMASE.
OS Lactococcus lactis.
OC Bacteria; Firmicutes; Lactobacilliales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1358;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MG1614;
RX MEDLINE=91197101; PubMed=1901704;
RA Sibakov M., Koivula T., von Wright A., Palva I.;
RT "Secretion of TEM beta-lactamase with signal sequences isolated from
the chromosome of Lactococcus lactis subsp. lactis.";
RL Appl. Environ. Microbiol. 57:341-348(1991).
DR EMBL; M63293; AAA64243.1; -.
FT NON_TER 72 72
SQ SEQUENCE 72 AA; 7428 MW; 16DFA92111A60D46 CRC64;

Query Match 2.0%; Score 8; DB 2; Length 72;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 75 SKAKASS 82
Db 26 SKAKASS 33
|||||

RESULT 21
Q8DB85 PRELIMINARY; PRT; 192 AA.
AC Q8DB85;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Anthranilate/para-aminobenzoate synthase component II.
GN V13065.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016807; AA011389.1; -.
RW Complete proteome.
SQ SEQUENCE 192 AA; 20563 MW; B537BF2BABA5946 CRC64;

Query Match 2.0%; Score 8; DB 16; Length 192;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 348 VPQTLTWT 355
Db 140 VPQTLTWT 147
|||||

RESULT 22
Q9TP09 PRELIMINARY; PRT; 227 AA.
AC Q9TP09;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE MHC class I antigen (Fragment).
GN MHC1UDA OR DABE-UDA.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20261668; PubMed=10799891;
RA Michaelova V., Murray B.W., Sultmann H., Klein J.;
RT "A contig map of the Mhc class I genomic region in the zebrafish
reveals ancient synteny.";
RL J. Immunol. 164:5296-5305(2000).
CC -I- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
IMMUNE SYSTEM (BY SIMILARITY).
CC -I- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
MICROGLOBULIN) (BY SIMILARITY).
CC EMBL; AF182155; AAF20178.1; -.
DR ZFIN; ZDB-GENE-010116-8; mhcluda.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR ProDom; PD000050; MHC_I; 1.
DR SMART; SM00407; IGC1; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
KW Glycoprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 227 227
SQ SEQUENCE 227 AA; 26616 MW; 22C1B6E02FF81D79 CRC64;

Query Match 2.0%; Score 8; DB 7; Length 227;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 351 TLTVTPEE 358
Db 215 TLTVTPEE 222
|||||

RESULT 23
Q8XRR4 PRELIMINARY; PRT; 248 AA.
AC Q8XRR4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)

```

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Probable transmembrane protein.
 GN RSP0767 OR RS01941.
 OS *Ralstonia solanacearum* (*Pseudomonas solanacearum*).
 OG Plasmid megaplasmid.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Ralstoniaceae; Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GMI1000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
 RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
 RA Siguer P., Thebault P., Whalen M., Wincker P., Levy M.,
 RA Weissenbach J., Boucher C.A.;
 RL "Genome sequence of the plant pathogen *Ralstonia solanacearum*."
 RL Nature 415:497-502(2002).
 DR EMBL; AL646080; CAD17918.1; -.
 KW Plasmid; Complete proteome.
 SQ SEQUENCE 248 AA; 24622 MW; 700F48E0F6B9A31B CRC64;

Query Match 2.0%; Score 8; DB 16; Length 248;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 AQAAPVA 104
 DB 94 AQAAPVA 101
 |||||

RESULT 24
 Q8ZBJO PRELIMINARY; PRT; 254 AA.
 AC Q8ZBJO;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Pyruvate dehydrogenase complex repressor.
 GN PDHR OR YPO3420.
 OS *Yersinia pestis*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OX NCBI_TaxID=632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CO-92 / Biovar Orientalis;
 RX MEDLINE=21470413; PubMed=11586360;
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
 RA Prentice M.B., Sebaiha M., James K.D., Churcher C., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdono-Tarraga A.M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
 RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
 RL "Genome sequence of *Yersinia pestis*, the causative agent of plague."
 RL Nature 413:523-527(2001).
 DR EMBL; AJ414157; CAC92650.1; -.
 DR InterPro; IPR000524; HTH_GNTR.
 DR Pfam; PF00392; gntr.1.
 DR PRINTS; PR00035; HTHGNTR.
 DR SMART; SM00345; HTH_GNTR; 1.
 DR PROSITE; PS00043; HTH_GNTR_FAMILY; 1.
 KW Complete proteome.
 SQ SEQUENCE 254 AA; 29408 MW; 24BACF565816CFE0 CRC64;

Query Match 2.0%; Score 8; DB 16; Length 254;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 REAIQRL 366
 DB 52 REAIQRL 59
 |||||

RESULT 25
 Q8DC23 PRELIMINARY; PRT; 255 AA.
 AC Q8DC23;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Pyruvate dehydrogenase complex repressor.
 GN VV11629.
 OS *Vibrio vulnificus*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=672;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CMCP6;
 RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
 RA Choy H.E.;
 RT "Complete genome sequence of *Vibrio vulnificus* CMCP6."
 RL Submitted (DEC-2002) to the EMBL/GenBank/DDRJ databases.
 DR EMBL; AE016802; AAC10048.1; -.
 KW Complete proteome.
 SQ SEQUENCE 255 AA; 28981 MW; 74965A33E5676FF6 CRC64;

Query Match 2.0%; Score 8; DB 16; Length 255;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 REAIQRL 366
 DB 52 REAIQRL 59
 |||||

RESULT 26
 Q9KPF3 PRELIMINARY; PRT; 256 AA.
 AC Q9KPF3;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Pyruvate dehydrogenase complex repressor.
 GN VC2415.
 OS *Vibrio cholerae*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=El Tor N16961 / Serotype O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Basso S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*."
 RL Nature 406:477-483(2000).
 CC -1- SIMILARITY: BELONGS TO THE GNTR FAMILY OF TRANSCRIPTIONAL REGULATORS.
 DR EMBL; AE004311; AAF95558.1; -.
 DR TIGR; VC2415; -.
 DR InterPro; IPR000524; HTH_Gntr.
 DR Pfam; PF00392; gntr.1.
 DR PRINTS; PR00035; HTHGNTR.
 DR SMART; SM00345; HTH_GNTR; 1.

DR PROSITE; PS00043; HTH_GNTR_FAMILY; 1.
 KW DNA-binding; Transcription regulation; Complete proteome.
 SQ SEQUENCE 256 AA; 29384 MW; C40833B79005A24B CRC64;

Query Match 2.0%; Score 8; DB 16; Length 256;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 REAIQRLE 366
 |||||
 DB 52 REAIQRLE 59

RESULT 27

Q52007 PRELIMINARY; PRT; 264 AA.
 AC Q52007;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Streptomycin-resistance protein.
 OS Vibrio fischeri.
 OG Plasmid pRL1063a.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=660;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21519270; PubMed=11607193;
 RA Wolk C.P., Cai Y., Panoff J.-M.,
 RT "Use of a transposon with luciferase as a reporter to identify
 RL environmentally responsive genes in a cyanobacterium.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:5355-5359(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Wolk C.P.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U55385; AA800446.1; -;
 DR InterPro; IPR006748; APH_6_hur.
 DR Pfam; PF04655; APH_6_hur; I.
 KW Plasmid.
 SQ SEQUENCE 264 AA; 28853 MW; 92AA3FDDF5BCF745 CRC64;

Query Match 2.0%; Score 8; DB 2; Length 264;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 APAASVAR 111
 |||||
 DB 132 APAASVAR 139

RESULT 28

O80422 PRELIMINARY; PRT; 272 AA.
 AC O80422;
 DT 01-NOV-1998 (TReMBLrel. 08, Created)
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
 DE Carbonic anhydrase.
 GN P0004A09.27.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Takano T., Liu S.;
 RT "Cloning of a gene for carbonic anhydrase in rice.";
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 clone: P0004A09.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB016283; BAA31953.1; -;
 DR EMBL; AP003607; BAE63789.1; -;
 DR Gramene; O80422; -;
 DR InterPro; IPR001765; Prok_COanhhd.
 DR Pfam; PF00484; Pro_CA; 1.
 DR PROSITE; PS00704; PROK_CO2_ANNHYDRASE_1; 1.
 DR PROSITE; PS00705; PROK_CO2_ANNHYDRASE_2; 1.
 SQ SEQUENCE 272 AA; 29117 MW; F414E718F846C5D2 CRC64;

Query Match 2.0%; Score 8; DB 10; Length 272;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 AAPVAPAA 107
 |||||
 DB 63 AAPVAPAA 70

RESULT 29

Q40628 PRELIMINARY; PRT; 273 AA.
 AC Q40628;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
 DE Carbonic anhydrase.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUR=Leaf;
 RA Suzuki S., Burnell J.N.;
 RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. IR64;
 RA Wang W., Zhu L.;
 RT "Isolation and characterization of a gene for carbonic anhydrase in
 RT rice.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U08404; AAA86943.1; -;
 DR EMBL; AF182806; AAD56038.1; -;
 DR Gramene; Q40628; -;
 DR InterPro; IPR001765; Prok_COanhhd.
 DR Pfam; PF00484; Pro_CA; 1.
 DR PROSITE; PS00704; PROK_CO2_ANNHYDRASE_1; 1.
 DR PROSITE; PS00705; PROK_CO2_ANNHYDRASE_2; 1.
 SQ SEQUENCE 273 AA; 29204 MW; 686EACB46DD9897A CRC64;

Query Match 2.0%; Score 8; DB 10; Length 273;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 AAPVAPAA 107
 |||||
 DB 64 AAPVAPAA 71

RESULT 30

Q8D1B5 PRELIMINARY; PRT; 278 AA.
 AC Q8D1B5;
 DT 01-MAR-2003 (TReMBLrel. 23, Created)
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)

DE Transcriptional regulator for pyruvate dehydrogenase complex.

GN PDHR OR Y0766.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Retherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM."
RL J. Bacteriol. 184:4601-4611(2002).
DR EMBL; AE013678; AA84353.1; --
KW Pyruvate.
SQ SEQUENCE 278 AA; 32020 MW; 9FC639740F8D0CD9 CRC64;

Query Match 2.0%; Score 8; DB 16; Length 278;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 359 REAIQRLR 366

Db 76 REAIQRLR 83

RESULT 31

ID Q9X2P4 PRELIMINARY; PRT; 289 AA.
AC Q9X2P4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Alkanal monooxygenase-like hypothetical protein.
GN AWLP.
OS Mycobacterium smegmatis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1772;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=mc2155;
RA Kim B.-K., Choi K.-P., Daniels L.;
RT "The gene for f420-dependent glucose-6-phosphate dehydrogenase (fgd)
RT and a conserved upstream gene coding for a beta-lactamase-like protein
RT in Mycobacterium species."
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF041061; AAD31326.1; --
DR InterPro; IPR002103; Bac_luciferase.
DR Pfam; PF00296; bac_luciferase; 1.
KW Hypothetical protein; Monooxygenase.
SQ SEQUENCE 289 AA; 31631 MW; 28776602C9BEFALE CRC64;

Query Match 2.0%; Score 8; DB 2; Length 289;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 267 VVFGAGSG 274

Db 114 VVFGAGSG 121

RESULT 32

O65002
ID O65002 PRELIMINARY; PRT; 300 AA.
AC O65002;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Isoflavone reductase homolog Bet v 6.0101 (Fragment).

GN BETV6.
OS Betula verrucosa (White birch) (Betula pendula).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Betula.
OX NCBI_TaxID=3505;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pollen;
RX MEDLINE=98178680; PubMed=9519865;
RA Vieths S., Frank E., Scheurer S., Meyer H.E., Hrazdina G.,
RA Hausteil D.;
RT "Characterization of a new IGB-binding 35-kDa protein from birch
RT pollen with cross-reacting homologues in various plant foods."
RL Scand. J. Immunol. 47:263-272(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pollen;
RX MEDLINE=20020186; PubMed=10550744;
RA Karamloo F., Schmitz N., Scheurer S., Foetisch K., Hoffmann A.,
RA Hausteil D., Vieths S.;
RT "Molecular cloning and characterization of a birch pollen minor
RT allergen, Bet v 5, belonging to a family of isoflavone reductase-
RT related proteins."
RL J. Allergy Clin. Immunol. 104:991-999(1999).
DR EMBL; AF135127; AAC05116.2; --
DR InterPro; IPR003866; Isoflav_reduct.
DR Pfam; PF02716; Isoflavone_redu; 1.
FT NON_TER 300 300
SQ SEQUENCE 300 AA; 33153 MW; D5B4DEC356E937C2 CRC64;

Query Match 2.0%; Score 8; DB 10; Length 300;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 TNFIEAS 18

Db 276 TNFIEAS 283

RESULT 33

O9FUW6 PRELIMINARY; PRT; 308 AA.
ID O9FUW6
AC O9FUW6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Allergenic isoflavone reductase-like protein Bet v 6.0102.
GN BETV6.0102.
OS Betula verrucosa (White birch) (Betula pendula).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Betula.
OX NCBI_TaxID=3505;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pollen;
RX MEDLINE=20020186; PubMed=10550744;
RA Karamloo F., Schmitz N., Scheurer S., Foetisch K., Hoffmann A.,
RA Hausteil D., Vieths S.;
RT "Molecular cloning and characterization of a birch pollen minor
RT allergen, Bet v 5, belonging to a family of isoflavone reductase-
RT related proteins."
RL J. Allergy Clin. Immunol. 104:991-999(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pollen;
RX Karamloo F., Wangorsch A., Hausteil D., Vieths S.;
RT "Characterization Bet v 6.0102, a high-IGE binding variant of the
RT birch pollen minor allergen, Bet v 6."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF282850; AAG22740.1; --

```

DR InterPro; IPR003866; Isoflav reduct.
DR Pfam; PF02716; Isoflavone reduct; 1.
SQ SEQUENCE 308 AA; 34130 MW; C14AB150751542BC CRC64;

Query Match      2.0%; Score 8; DB 10; Length 308;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      11 TNFEIEAS 18
Db      276 TNFEIEAS 283

RESULT 34
Q8HWF2 PRELIMINARY; PRT; 322 AA.
AC Q8HWF2;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE SI:G2179B20.4 (Major histocompatibility complex class I UDA gene)
DE (Fragment).
GN MHC1UDA.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Hammond S.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL672151; CAD58766.1; -.
FT NON_TER 322 322
SQ SEQUENCE 322 AA; 37121 MW; BD1C4B175DDA90DD CRC64;

Query Match      2.0%; Score 8; DB 7; Length 322;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      351 TLTVTPEE 358
Db      257 TLTVTPEE 264

RESULT 35
Q9GJKO PRELIMINARY; PRT; 339 AA.
AC Q9GJKO;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE MHC class I heavy chain (Fragment).
GN SATR-UBA.
OS Salmo trutta (Brown trout).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8032;
RN [1]
RP SEQUENCE FROM N.A.
RA Tissue-Spleen.
RC TISSUE=Spleen.
EX MEDLINE=21142794; PubMed=11207285;
RA Shum B.P., Guethlein L., Flodin L.R., Adkison M.A., Hedrick R.P.,
RA Nehring R.B., Stet R.J.M., Secombes C., Parham P.;
RT "Modes of Salmonid MHC Class I and II Evolution Differ from the
RT Primate Paradigm."
RL J. Immunol. 166:3297-3308 (2001).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC IMMUNE SYSTEM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN) (BY SIMILARITY).
DR EMBL; AF296376; AAG02522.1; -.

```

```

DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR PRODOM; PD000050; MHC_I; 1.
DR SMART; SM00407; Igel; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Glycoprotein; Transmembrane.
FT NON_TER 339 339
SQ SEQUENCE 339 AA; 37983 MW; 67730F33F1E4981D CRC64;

Query Match      2.0%; Score 8; DB 7; Length 339;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      352 LTVTPEER 359
Db      250 LTVTPEER 257

RESULT 36
P91201 PRELIMINARY; PRT; 354 AA.
AC P91201;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical 41.9 kDa protein.
GN EGAP9.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA None;
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Geisel C.;
RT "The sequence of C. elegans cosmid EGAP9.";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U80026; AAC25845.2; -.
DR WormPep; EGAP9.3; CE27912.
DR InterPro; IPR002516; Glyco trans 11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Hypothetical protein.
SQ SEQUENCE 354 AA; 41879 MW; E0F60E58CAA1BCDA CRC64;

Query Match      2.0%; Score 8; DB 5; Length 354;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      193 RAIDYLYS 200
Db      230 RAIDYLYS 237

RESULT 37

```

Q8PNM4	PRELIMINARY;	PRT;	355 AA.
ID	Q8PNM4		
AC	Q8PNM4		
DT	01-OCT-2002 (TrEMBLrel. 22, Created)		
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)		
DE	Hypoethetical protein XAC1030.		
GN	XAC1030.		
OS	Xanthomonas axonopodis (pv. citri).		
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;		
OC	Xanthomonadaceae; Xanthomonas.		
OX	NCBI_TaxID=92829;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=306 / ATCC 13902 / XV 101;		
RX	MEDLINE=22022145; PubMed=12024217;		
RA	da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,		
RA	Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,		
RA	Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,		
RA	Cararotte G., Cannavan F., Cardoso J., Chamberg F., Ciapina L.P.,		
RA	Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,		
RA	Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,		
RA	Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.		
RA	Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,		
RA	Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,		
RA	Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,		
RA	Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,		
RA	Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,		
RA	Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,		
RA	Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,		
RA	Setubal J.C., Kitajima J.P.;		
RT	"Comparison of the genomes of two Xanthomonas pathogens with differing		
RT	host specificities."		
RL	Nature 417:459-463(2002).		
DR	EWBL: AE011732; AM35913.1; "		
DR	InterPro; IPRO02965; P-rich extensin.		
DR	PRINTS; PR01217; PRICHEXTENS.		
KW	Hypoethetical protein; Complete proteome.		
SQ	SEQUENCE 355 AA; 35642 MW; 195C067A70E75107 CRC64;		
Query Match	2.0%; Score 8; DB 16; Length 355;		
Best Local Similarity	100.0%; Pred. No. 53;		
Matches	8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	101 APVAPAAAS 108		
D6	270 APVAPAAAS 277		
RESULT 38			
Q8Y277	PRELIMINARY;	PRT;	356 AA.
ID	Q8Y277		
AC	Q8Y277		
DT	01-MAR-2002 (TrEMBLrel. 20, Created)		
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)		
DE	Hypoethetical transmembrane protein RSC0459.		
GN	RSC0459 OR RS04440.		
OS	Ralstonia solanacearum (Pseudomonas solanacearum).		
OC	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;		
OC	Ralstoniaceae; Ralstonia.		
OX	NCBI_TaxID=305;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=GM11000;		
RX	MEDLINE=21681879; PubMed=11823852;		
RA	'alanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,		
RA	Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,		
RA	Chandler M., Choise N., Claudel-Renard C., Cunncan S., Demange N.,		
RA	Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,		
RA	Stguier P., Thebaud P., Whalen M., Wincker P., Levy M.,		
RA	Weissenbach J., Boucher C.A.;		
RT	"Genome sequence of the plant pathogen Ralstonia solanacearum";		

```

RN  SEQUENCE FROM N.A.
RP  STRAIN=21G1;
RC  MEDLINE=21844508; PubMed=11855723; Quesada A.;
RA  Genez I., Mehan P., Fernandez E.,
RT  "NADP-malate dehydrogenase from Chlamydomonas: prediction of new
RL  structural determinants for redox regulation by homology modelling.";
RC  Plant Mol. Biol. 48:211-221(2002).
CC  -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC  -1- SIMILARITY: BELONGS TO THE LDH FAMILY. MDH SUBFAMILY.
DR  EMBL; AJ272781; CAC19083.2; -.
DR  HSPF; P17606; 7MDH.
DR  InterPro; IPR001236; ldh.
DR  InterPro; IPR001252; Mdh.
DR  Pfam; PF00056; ldh; 1.
DR  Pfam; PF02866; ldh_C; 1.
DR  ProDom; PD003052; Mdh; 1.
DR  PROSITE; PS00068; MDH; 1.
DR  KX  Oxidoreductase; NADP.
SQ  SEQUENCE 415 AA; 44852 MW; 7B32E23DA26D53C8 CRC64;

Query Match          2.0%; Score 8; DB 10; Length 415;
Best Local Similarity 100.0%; Pred.No. 61;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  100 AAPVAPAA 107
DB  |||||
    32 AAPVAPAA 39

RESULT 41
ID  Q9DBZ2 PRELIMINARY; PRT; 425 AA.
AC  Q9DBZ2;
DT  01-JUN-2001 (TrEMBLrel. 17, Created)
DT  01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT  01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE  Speckle-type POZ protein.
GN  SPOB.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Lung;
RX  MEDLINE=21085660; PubMed=11217851;
RA  Kawai J., Shinagawa A., Shibata K., Konno H., Adachi J., Fukuda S.,
RA  Arakawa T., Hara A., Fukunishi Y., Kiyosawa H., Kondo S., Yamanaka I.,
RA  Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA  Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA  Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA  Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA  Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA  Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA  Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA  Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA  Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA  Guatinchich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA  Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA  Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA  Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
RA  Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA  Wynshaw-Boris A., Yoshida K., Haegawa Y., Kawaji H., Kohtsuki S.,
RA  Hayashizaki Y.;
RT  "Functional annotation of a full-length mouse cDNA collection.";
RL  Nature 409:685-690(2001).
DR  EMBL; AK004669; BAB23458.1; -.
DR  MGB; MGI:1343085; Spop.
DR  InterPro; IPR000210; BTB_POZ.
DR  InterPro; IPR002083; MATH.
DR  Pfam; PF00651; BTB; 1.
DR  Pfam; PF00917; MATH; 1.

```

```

DR  SMART; SM00225; BTB; 1.
DR  SMART; SM00061; MATH; 1.
DR  PROSITE; PS00097; BTB; 1.
SQ  SEQUENCE 425 AA; 47514 MW; 7B6716666C3D884B CRC64;

Query Match          2.0%; Score 8; DB 11; Length 425;
Best Local Similarity 100.0%; Pred.No. 63;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  135 AATVAATD 142
DB  |||||
    390 AATVAATD 397

RESULT 42
ID  Q9HP89 PRELIMINARY; PRT; 462 AA.
AC  Q9HP89;
DT  01-MAR-2001 (TrEMBLrel. 16, Created)
DT  01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT  01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE  Photolyase/cryptochrome.
GN  PHR1 OR VNG17546.
OS  Halobacterium sp. (strain NRC-1).
OC  Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC  Halobacteriaceae; Halobacterium.
OX  NCBI_TaxID=64091;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=20504483; PubMed=11016950;
RA  Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA  Shukla H.D., Laaky S.R., Baliga N.S., Thorsson V., Sbrocna J.,
RA  Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA  Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA  Maddocks D.G., Jablonaki P.E., Krebs M.P., Angevine C.M., Dale H.,
RA  Ikenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA  Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA  Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT  "Genome sequence of Halobacterium species NRC-1";
RL  Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR  EMBL; AE005079; AAG19981.1; -.
DR  HSPF; P05327; IQNF.
DR  InterPro; IPR006049; DNA_photolyase.
DR  InterPro; IPR002081; DNA_photolyase_1.
DR  InterPro; IPR006051; DNA_photolyase_C.
DR  InterPro; IPR006050; DNA_photolyase_N.
DR  InterPro; IPR005101; FAD_binding_7.
DR  InterPro; IPR00209; Peptidase_S8.
DR  Pfam; PF00875; DNA_photolyase; 1.
DR  Pfam; PF03441; FAD_binding_7; 1.
DR  PRINTS; PR00147; DNAPHOTLYASE.
DR  ProDom; PD004390; DNA_photolyase_C; 1.
DR  PROSITE; PS00394; DNA_PHOTOLYASES_1; 1.
DR  PROSITE; PS00136; SUBTILASE_ASP; 1.
KW  Lyase; Complete proteome.
SQ  SEQUENCE 462 AA; 51430 MW; 2CB9EE1CFA3976AB CRC64;

Query Match          2.0%; Score 8; DB 17; Length 462;
Best Local Similarity 100.0%; Pred.No. 68;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  181 VRALRAY 188
DB  |||||
    53 VRALRAY 60

RESULT 43
ID  Q9RK92 PRELIMINARY; PRT; 466 AA.
AC  Q9RK92;
DT  01-MAY-2000 (TrEMBLrel. 13, Created)
DT  01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT  01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

```



```

DE Hypothetical protein SC00270.
GN SCO0270 OR SCF1.12
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Baten A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RL coelicolor A3(2)." (2002).
RL Nature 417:141-147(2002).
DR EMBL; AL939104; CAB5532.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 466 AA; 49047 MW; 567F8227E4261096 CRC64;

Query Match 2.0%; Score 8; DB 16; Length 466;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 PAATVAAT 141
DB 52 PAATVAAT 59
|||||

RESULT 44
Q8RVT6 PRELIMINARY; PRT; 467 AA.
AC Q8RVT6;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DE 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE PSAPV2.
GN PSAPV2.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. midoriisui; TISSUE=Etiolated stem;
RA Kawanara T., Miura A., Kiba A., Toyoda K., Ichinose Y., Shiraiishi T.;
RT "Characterization of Pea Cell Wall-bound Apyrase."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB071370; BAB85978.1; -.
DR InterPro; IPR000407; GDAL_CD39_NTPase.
DR Pfam; PF01150; GDAL_CD39; 1.
DR PROSITE; PS01238; GDAL CD39 NTPASE; 1.
SQ SEQUENCE 467 AA; 50964 MW; 94167066C46E0D3C CRC64;

Query Match 2.0%; Score 8; DB 10; Length 467;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 KASSSSGAS 85
DB 307 KASSSSGAS 314
|||||

RESULT 45
Q92KE5 PRELIMINARY; PRT; 473 AA.
ID Q92KE5
AC Q92KE5;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical protein RO1411.
GN RO1411 OR SMC00996.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Porcette D., Puehler A., Fumelle B., Rampsperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL; AL591787; CAC45990.1; -.
DR InterPro; IPR001967; Ala/AlaCBptase1.
DR Pfam; PF00768; Peptidase S11; 1.
DR PRINTS; PR00725; DADACBPTASE1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 473 AA; 49918 MW; 15044040A2B6725 CRC64;

Query Match 2.0%; Score 8; DB 16; Length 473;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 VATAETAP 125
DB 372 VATAETAP 379
|||||

RESULT 46
Q8SOB0 PRELIMINARY; PRT; 480 AA.
ID Q8SOB0;
AC Q8SOB0;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Putative RNA helicase.
GN P0470A12.6 OR P0456E05.28.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartioideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
RT clone:P0470A12."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
RT clone:P0456E05."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
DR EMBL; AP003436; BAB90281.1; -.
DR EMBL; AP003416; BAB92627.1; -.
DR Gramene; Q8SOB0; -.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR000629; DEAD_box.
DR InterPro; IPR001650; Helicase_C.

```

DR Pfam; PF00270; DEAD; 1.
 DR Pfam; PF00271; helicase_C; 1.
 DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
 KW ATP-binding; Helicase; Hydrolase.
 SQ SEQUENCE 480 AA; 52191 MW; DB67DB68D6A73994 CRC64;

Query Match 2.0%; Score 8; DB 10; Length 480;
 Best Local Similarity 100.0%; Pred. No. 70;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 99 PAAPVAPA 106
 |||||
 Db 46 PAAPVAPA 53

RESULT 47
 Q8Y1V5 PRELIMINARY; PRT; 514 AA.
 AC Q8Y1V5;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Probable D-alanyl-D-alanine carboxypeptidase signal peptide
 DE protein.
 GN RSC0584 OR RS04879.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Ralstoniaceae; Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GM1000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Arianoutat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Arianoutat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
 RA Chandler M., Choise N., Claudel-Renard C., Cunac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
 RA Siguer P., Thebaud P., Whalen M., Wincker P., Levy M.,
 RA Weissenbach J., Boucher C.A.;
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum";
 RL Nature 415:497-502(2002).
 DR EMBL; AL646060; CAD14114.1; -;
 DR InterPro; IPR000867; Peptidase_S13.
 DR Pfam; PF02113; Peptidase_S13; 1.
 DR PRINTS; PR00922; DADACBPTASE3.
 DR TIGRFAMs; TIGR00666; PB4; 1.
 KW Complete proteome.
 SQ SEQUENCE 514 AA; 54615 MW; 2356A4830B15EED8 CRC64;

Query Match 2.0%; Score 8; DB 16; Length 514;
 Best Local Similarity 100.0%; Pred. No. 75;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 132 AAPATVA 139
 |||||
 Db 42 AAPATVA 49

RESULT 48
 P72770 PRELIMINARY; PRT; 566 AA.
 AC P72770;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE High affinity sulfate transporter.
 GN SLR1776.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;

RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hirose M., Sugita M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
 RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 Synechocystis sp. strain PCC6803. II. Sequence determination of the
 entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 DR EMBL; D90900; BAA16785.1; -;
 DR InterPro; IPR002645; STAS.
 DR InterPro; IPR001902; Sulph_transpt.
 DR Pfam; PF01740; STAS; 1.
 DR Pfam; PF00916; Sulfate transp; 1.
 DR TIGRFAMs; TIGR00815; sulp; 1.
 DR PROSITE; PS50801; STAS; 1.
 KW Complete proteome.
 SQ SEQUENCE 566 AA; 60928 MW; 11BA8DA1B0FC2FE6 CRC64;

Query Match 2.0%; Score 8; DB 16; Length 566;
 Best Local Similarity 100.0%; Pred. No. 82;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 257 GVPSSGSN 264
 |||||
 Db 155 GVPSSGSN 162

RESULT 49

Q9VUD3 PRELIMINARY; PRT; 571 AA.
 AC Q9VUD3;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE CG32139 protein (GH07353p).
 GN SOX21B OR SOX-LIKE OR CG6419 OR CG13483 OR CG32139.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Rogers J., Finkbeiner S., Miller L., Gilman J.,
 April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 BAILEW R.M., BAEU A., BAXENDALE J., BAYKATROGLU L., BEASLEY E.M.,
 BEESON K.Y., BENOS P.V., BERMAN B.P., BHANDARI D., BOLSHAKOV S.,
 BORKOVA D., BOTCHAN M.R., BOUCK J., BROKSTEIN P., BROTTIER P.,
 BURTIS K.C., BUSAM D.A., BUTLER H., CADIEU E., CENTER A., CHANDRA I.,
 CHERRY J.M., CRAWLEY S., DAHLKE C., DAVENPORT L.B., DAVIES P.,
 DE PABLOS B., DELCHER A., DENG Z., MAYS A.D., DEW I., DIETZ S.M.,
 DODSON K., DOUP L.E., DOWNES M., DUGAN-ROCHA S., DUNKOV B.C., DUNN P.,
 DURBIN K.J., EVANGELISTA C.C., FERTAZ C., FERREIRA S., FLEISCHMANN W.,
 FOELLER C., GABRIELIAN A.E., GARG N.S., GELBART W.M., GLASSER K.,
 GLUCK A., GONG F., GORELL J.H., GU Z., GUAN P., HARRIS M.,
 HARRIS N.L., HARVEY D., HEIMAN T.J., HERNANDEZ J.R., HOWACK J.,
 HOUSTIN D., HOUSTON K.A., HOWLAND T.J., WEI M.-H., IBEGWAN J.,
 RA JALALI M., KALUSH F., KARPEN G.H., KE Z., KENNISON J.A., KATCHUM K.A.,
 KIMMEL B.E., KODIRA C.D., KRAFT C., KRAVITZ S., KULP D., LAI Z.,
 RA LAOKO P., LEI Y., LEVITSKY A.C., LI J., LI Z., LIANG Y., LIN X.,
 RA LIU X., MATTEI B., MCINTOSH T.C., MCLEOD M.P., MCPHERSON D.,
 RA MERKULOV G., MILSHINA N.V., MURPHY B., MORRIS J., MOSHREFI A.,
 RA MOUNT S.M., MOY M., MURPHY B., MURPHY L., MURPHY D.M., NELSON D.L.,
 RA NELSON D.R., NELSON K.A., NIXON K., NUSKERN D.R., PACLEB J.M.,
 PALAZZOLO M., PITTMAN G.S., PAN S., POLLARD J., PURI V., REESE M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195 (2000).
RN [2]
RN SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banton J., An H., Baldwin D., Bantz J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of *Drosophila melanogaster* genome.";
RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of *Drosophila melanogaster* genome.";
RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RN SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RN SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RN SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarino H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AS003535; AAF49754.2; -;
DR EMBL; AY058313; AAL13542.1; -;
DR HSSP; Q05066; 1HRX.
DR FlyBase; FBgn0042630; Sox21b.
DR InterPro; IPR000910; HMG_12_box.
DR Pfam; PF00505; HMG_box; 1.
DR SMART; SM00398; HMG; 1.
SQ SEQUENCE 571 AA; 60909 MW; 123F3C1C3AACAF29 CRC64;
Query Match 2.0%; Score 8; DB 5; Length 571;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 341 LGQLAAAV 348
DB 33 LGQLAAAV 40

RESULT 50

O85688 PRELIMINARY; PRT; 595 AA.
ID 085688
AC 085688;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Flagellar hook length determination protein homolog.
GN FLHG.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC NCBI_TaxID=155892;
OX [1]
RN SEQUENCE FROM N.A.
RC STRAIN=CB15;
RX MEDLINE=89255087; PubMed=2470725;
RA Mullin D.A., Newton A.;
RT "Ntr-like promoters and upstream regulatory sequence ftr are required
for transcription of a developmentally regulated Caulobacter
crescentus flagellar gene.";
RT J. Bacteriol. 171:3218-3227(1989).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=CB15;
RA Mullin D.A., Mullin A.H., Ohta N., Newton A.;
RT "Organization and expression of Caulobacter crescentus genes needed
for assembly and function of the flagellar hook.";
RL MOL. Gen. Genet. 265:445-454(2001).
DR EMBL; AF072135; AAC33327.1; -;
DR InterPro; IPR000104; Antifreeze_1.
DR PRINTS; PR00308; ANTIFREEZE1.
SQ SEQUENCE 595 AA; 58056 MW; 218BE9BBBA3C93E CRC64;
Query Match 2.0%; Score 8; DB 2; Length 595;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 51

P72058 PRELIMINARY; PRT; 643 AA.
ID P72058
AC P72058;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein RV3792.
GN RV3792 OR MTCY13D12.26.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badoock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
RL Nature 393:537-544(1998).
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).

```

CC      -!- SIMILARITY: IN THE C-TERMINAL REGION, TO M.SMEGMATIS KATG.
DR      EMBL; Z80343; CAB02471.1; -.
DR      Hyperculist; RV3792; -.
KW      Hypoetical protein; Transmembrane; Complete proteome.
FT      TRANSMEM 25 45
FT      TRANSMEM 66 86
FT      TRANSMEM 97 117
FT      TRANSMEM 151 171
FT      TRANSMEM 180 200
FT      TRANSMEM 205 225
FT      TRANSMEM 229 249
FT      TRANSMEM 271 291
FT      TRANSMEM 309 329
FT      TRANSMEM 355 375
FT      TRANSMEM 384 404
FT      TRANSMEM 420 440
FT      TRANSMEM 445 465
FT      TRANSMEM 517 537
FT      TRANSMEM 616 636
SQ      SEQUENCE 643 AA; 69515 MW; 7C58972B085D9EDA CRC64;

Query Match      2.0%; Score 8; DB 16; Length 643;
Best Local Similarity 100.0%; Pred. No. 93;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      341 LGQLAAAV 348
Db      29 LGQLAAAV 36

RESULT 52
Q8YGD6      PRELIMINARY; PRT; 653 AA.
AC      Q8YGD6;
DT      01-MAR-2002 (TrEMBLrel. 20, Created)
DT      01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT      01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE      TRANSPOSASE.
GN      BMEI1223.
OS      Brucella melitensis.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC      Brucellaceae; Brucella.
OX      NCBI_TaxID=29459;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=16M / ATCC 23456 / Biotype 1;
RX      MEDLINE=20020109; PubMed=1175668;
RA      DelVecchio V.G., Kapratral V., Redkar R.J., Patra G., Mujar C., Los T.,
RA      Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
RA      Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
RA      Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA      Haselkorn R., Kyripides N., Overbeek R.;
RT      "The genome sequence of the facultative intracellular pathogen
RT      Brucella melitensis.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:443-448 (2002).
DR      EMBL; AE009561; AAL52404.1; -.
DR      InterPro; IPR001584; Rve.
DR      Pfam; PF00665; Rve; 1.
KW      Complete proteome.
SQ      SEQUENCE 653 AA; 73916 MW; CB1A023F13D95954 CRC64;

Query Match      2.0%; Score 8; DB 16; Length 653;
Best Local Similarity 100.0%; Pred. No. 94;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      273 SGALDALR 280
Db      255 SGALDALR 262

RESULT 53
Q9ABQ3      PRELIMINARY; PRT; 687 AA.
ID      Q9ABQ3

```

```

AC      Q9ABQ3;
DT      01-JUN-2001 (TrEMBLrel. 17, Created)
DT      01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      General secretion pathway protein D.
GN      CC0173.
OS      Caulobacter crescentus.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC      Caulobacteraceae; Caulobacter.
OX      NCBI_TaxID=155892;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 19089 / CB15;
RX      MEDLINE=21173698; PubMed=11259647;
RA      Nierman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA      Bisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA      Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA      DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA      Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA      Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA      Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT      "Complete genome sequence of Caulobacter crescentus.";
RL      Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141 (2001).
DR      EMBL; AE005692; AAK22160.1; -.
DR      TIGR; CC0173; -.
DR      InterPro; IPR001775; Bac_GSPD.
DR      InterPro; IPR004846; GSP1/IIIProtein.
DR      InterPro; IPR005644; NoliW-like.
DR      Pfam; PF00263; GSP11_III; 1.
DR      Pfam; PF03958; GSP11_III_N; 3.
DR      PRINTS; PR00811; BCTERIALGSPD.
KW      Complete proteome.
SQ      SEQUENCE 687 AA; 71703 MW; 1EEDBB7E4B11B563 CRC64;

Query Match      2.0%; Score 8; DB 16; Length 687;
Best Local Similarity 100.0%; Pred. No. 99;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      83 GASTATTA 90
Db      307 GASTATTA 314

RESULT 54
Q8CI48      PRELIMINARY; PRT; 695 AA.
ID      Q8CI48
AC      Q8CI48;
DT      01-MAR-2003 (TrEMBLrel. 23, Created)
DT      01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      Similar to expressed sequence A1118201 (Fragment).
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=FVB/N.
RA      Strauberg R.;
RL      Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL; BC037480; AAH37480.1; -.
FT      NON TER 1 1
SQ      SEQUENCE 695 AA; 73412 MW; 46791F8FFA57E5DE CRC64;

Query Match      2.0%; Score 8; DB 11; Length 695;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      134 PAATVAAT 141
Db      7 PAATVAAT 14

```

RESULT 55

Q8G1I6 PRELIMINARY; PRT; 704 AA.
 AC Q8G1I6;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Transposase, putative.
 GN BR0729.
 OS Brucella suis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29461;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1330 / Biovar 1;
 RX MEDLINE=22247741; PubMed=12271122;
 RA Paulsen I.T., Seehadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
 RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
 RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
 RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,
 RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
 RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.,
 RT "The Brucella suis genome reveals fundamental similarities between
 RT animal and plant pathogens and symbionts."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
 DR EMBL; A5014378; AAN29658.1; -.
 DR TIGR; BR0729; -.
 KW Complete proteome.
 SQ SEQUENCE 704 AA; 79768 MW; 6A47F42C372995AD CRC64;

Query Match 2.0%; Score 8; DB 16; Length 704;
 Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

273 SGALDAR 280
 306 SGALDAR 313
 |||||

RESULT 56

ID Q8VIT6 PRELIMINARY; PRT; 757 AA.
 AC Q8VIT6;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein MT3899.1.
 GN MT3899.1.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; A5007183; AAK48265.1; -.
 DR TIGR; MT3899; -.
 KW Hypothetical protein.
 SQ SEQUENCE 757 AA; 82352 MW; 30210944EF4BEAAE CRC64;

Query Match 2.0%; Score 8; DB 16; Length 757;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 LGQLAAAV 348
 |||||
 Db 29 LGQLAAAV 36

RESULT 57

ID Q8K1S4 PRELIMINARY; PRT; 898 AA.
 AC Q8K1S4;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Netrin receptor Unc5h1.
 GN UNC5H1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Engelkamp D.;
 RT "Cloning of three mouse unc-5 genes and their expression patterns at
 RT mid-gestation."
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ487852; CAD32250.1; -.
 DR MGI; MGI:894582; Unc5h1.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003006; IG MHC.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR000906; ZU5.
 DR Pfam; PF00531; death; 1.
 DR Pfam; PF00047; Ig; 1.
 DR Pfam; PF00090; tsp.1; 2.
 DR Pfam; PF00791; ZU5; 1.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00209; TSP1; 2.
 DR PROSITE; PS00835; IG LIKE; 1.
 DR PROSITE; PS50092; TSP1; 2.
 KW Receptor.
 SQ SEQUENCE 898 AA; 98856 MW; 59F04BA2E196C1DB CRC64;

Query Match 2.0%; Score 8; DB 11; Length 898;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 LGQLAAAV 348
 |||||
 Db 873 LGQLAAAV 880

RESULT 58

ID Q08721 PRELIMINARY; PRT; 898 AA.
 AC Q08721;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Transmembrane receptor UNC5H1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain, and Ventral spinal cord;
 RX MEDLINE=97271897; PubMed=9126742;
 RA Leonardo E.D., Hinc L., Masu M., Keino-Masu K., Ackerman S.L.,
 RA Tessier-Lavigne M.;
 RT "Vertebrate homologues of C. elegans UNC-5 are candidate netrin

```

RT  receptors.;
RL  Nature 386:833-838(1997).
DR  EMBL; U87305; AAB57678.1; -.
DR  InterPro; IPR000488; Death.
DR  InterPro; IPR003599; IG.
DR  InterPro; IPR007110; IG-like.
DR  InterPro; IPR003006; IG-MHC.
DR  InterPro; IPR000884; TSF1.
DR  InterPro; IPR000906; ZU5.
DR  Pfam; PF00531; death; 1.
DR  Pfam; PF00047; ig; 1.
DR  Pfam; PF00090; Lep_1; 2.
DR  Pfam; PF00791; ZU5_1; 1.
DR  SMART; SM00005; DEATH; 1.
DR  SMART; SM00409; IG; 1.
DR  SMART; SM00209; TSF1; 2.
DR  SMART; SM00218; ZU5; 1.
DR  PROSITE; PS0835; IG-LIKE; 1.
DR  PROSITE; PS0092; TSF1; 2.
DR  Receptor.
SQ  SEQUENCE 898 AA; 98840 MW; 7A3CBCB9E7ACA135 CRC64;

Query Match 2.0%; Score 8; DB 11; Length 898;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 341 LGQLAAAV 348
Db 873 LGQLAAAV 880

RESULT 59
Q8CJKB PRELIMINARY; PRT; 908 AA.
AC Q8CJKB
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Putative transcriptional regulator.
GN SC07173 OR SC8A11.01 OR SC9A4.35.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939130; CAD55526.1; -.
DR Complete proteome.
SQ SEQUENCE 908 AA; 98458 MW; C3D5B12D49623331 CRC64;

Query Match 2.0%; Score 8; DB 16; Length 908;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 342 GQLAAAVP 349
Db 375 GQLAAAVP 382

RT  receptors.;
RL  Nature 386:833-838(1997).
DR  EMBL; U87305; AAB57678.1; -.
DR  InterPro; IPR000488; Death.
DR  InterPro; IPR003599; IG.
DR  InterPro; IPR007110; IG-like.
DR  InterPro; IPR003006; IG-MHC.
DR  InterPro; IPR000884; TSF1.
DR  InterPro; IPR000906; ZU5.
DR  Pfam; PF00531; death; 1.
DR  Pfam; PF00047; ig; 1.
DR  Pfam; PF00090; Lep_1; 2.
DR  Pfam; PF00791; ZU5_1; 1.
DR  SMART; SM00005; DEATH; 1.
DR  SMART; SM00409; IG; 1.
DR  SMART; SM00209; TSF1; 2.
DR  SMART; SM00218; ZU5; 1.
DR  PROSITE; PS0835; IG-LIKE; 1.
DR  PROSITE; PS0092; TSF1; 2.
DR  Receptor.
SQ  SEQUENCE 898 AA; 98840 MW; 7A3CBCB9E7ACA135 CRC64;

Query Match 2.0%; Score 8; DB 11; Length 898;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 341 LGQLAAAV 348
Db 873 LGQLAAAV 880

RESULT 59
Q8CJKB PRELIMINARY; PRT; 908 AA.
AC Q8CJKB
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Putative transcriptional regulator.
GN SC07173 OR SC8A11.01 OR SC9A4.35.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939130; CAD55526.1; -.
DR Complete proteome.
SQ SEQUENCE 908 AA; 98458 MW; C3D5B12D49623331 CRC64;

Query Match 2.0%; Score 8; DB 16; Length 908;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 342 GQLAAAVP 349
Db 375 GQLAAAVP 382

```

```

RESULT 60
Q922S3 PRELIMINARY; PRT; 944 AA.
AC Q922S3
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DB Unknown (Protein for IMAGE:3594992) (Fragment).
GN A118201.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 4 PDZ/DHR DOMAINS.
DR EMBL; BC006859; AA06859.1; -.
DR MGI; MGI:2145950; A118201.
DR InterPro; IPR001478; PDZ.
DR Pfam; PF00595; PDZ; 4.
DR SMART; SM00228; PDZ; 4.
DR PROSITE; PS0106; PDZ; 4.
FT NON TER 1 1
SQ SEQUENCE 944 AA; 99327 MW; 89CE11BA859B55E4 CRC64;

Query Match 2.0%; Score 8; DB 11; Length 944;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 134 PAATVAAT 141
Db 256 PAATVAAT 263

RESULT 61
Q9X7M2 PRELIMINARY; PRT; 1327 AA.
AC Q9X7M2
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE CTORF1365, partial (Fragment).
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COL;
RX MEDLINE=99265121; PubMed=10332717;
RA Wu S., de Lencastre H.;
RT "Mrp-a new auxiliary gene essential for optimal expression of
RT methicillin resistance in Staphylococcus aureus.";
RL Microb. Drug Resist. 5:9-18(1999).
DR EMBL; Y09928; CAA71062.1; -.
FT NON TER 1 1
FT NON TER 1327 1327
SQ SEQUENCE 1327 AA; 140517 MW; B90F2085E800586D CRC64;

Query Match 2.0%; Score 8; DB 2; Length 1327;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 97 AQAAPVA 104
Db 122 AQAAPVA 129

RESULT 62
Q8V111 PRELIMINARY; PRT; 1665 AA.
AC Q8V111

```

DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE PDZ-domain protein scribble.
GN A118201.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Cerebellum;
RA Mattock K.L., Kirschner C.;
RT "Molecular cloning of mouse Scribble cDNA."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 4 PDZ/DHR DOMAINS.
DR EMBL; AF441233; AAL32469.1; -;
DR MGD; MGI:2145950; A118201.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR_TYP.
DR InterPro; IPR001478; PDZ.
DR Pfam; PF00560; LRR; 13.
DR Pfam; PF00595; PDZ; 4.
DR SMART; SM00369; LRR_TYP; 2.
DR SMART; SM00228; PDZ; 4.
DR PROSITE; PS05056; LRR_TYPICAL; 2.
DR PROSITE; PS0106; PDZ_4.
SQ SEQUENCE 1665 AA; 179749 MW; B1272F4ECBF010E CRC64;

Query Match 2.0%; Score 8; DB 11; Length 1665;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 134 PAATVAAT 141
Db 952 PAATVAAT 959

RESULT 63
Q9LCJ9 PRELIMINARY; PRT; 1795 AA.
AC Q9LCJ9;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Truncated FmtB.
GN FMTB OR TRUNCATED FMTB OR MW2087.
OS Staphylococcus aureus, and
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280, 196620;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KSA8;
RX MEDLINE=20348625; PubMed=10896508;
RA Komatsuzaawa H., Ohta K., Sugai M., Fujiwara T., Glanzmann P.,
RA Berger-Bachi B., Suganaka H.;
RT "Fn551-mediated insertional inactivation of the fmbB gene encoding a
cell wall-associated protein abolishes methicillin resistance in
Staphylococcus aureus."
RT Staphylococcus aureus."
RL J. Antimicrob. Chemother. 45:421-431(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
acquired MRSA."
RL Lancet 359:1819-1827(2002).
DR EMBL; AB015223; BAA93430.1; -;
DR EMBL; AP004829; BAB95952.1; -;

DR InterPro; IPR00890; Acetate kin.
DR InterPro; IPR005877; Gpos_Ysirk.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
DR PROSITE; PS01075; ACETATE_KINASE_1; 1.
KW Complete proteome.
SQ SEQUENCE 1795 AA; 191008 MW; 6CB77CCFCB33D350 CRC64;

Query Match 2.0%; Score 8; DB 16; Length 1795;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 97 AQPAAPVA 104
Db 164 AQPAAPVA 171

RESULT 64
Q9RL69 PRELIMINARY; PRT; 2478 AA.
AC Q9RL69;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE MRP protein.
GN MRP.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COL;
RX MEDLINE=97302526; PubMed=9158773;
RA Wu S., de Lencastre H., Sali A., Tomasz A.;
RT "A phosphoglucosylase-like gene essential for the optimal expression
of methicillin resistance in Staphylococcus aureus: molecular cloning
and DNA sequencing."
RT Microb. Drug Resist. 2:277-286(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COL;
RX MEDLINE=97431478; PubMed=9286983;
RA Jolly L., Wu S.W., Van Heijenoort J., de Lencastre H.,
RA Mengin-Lecreulx D., Tomasz A.;
RT "The femR315 gene from Staphylococcus aureus, the interruption of
which results in reduced methicillin resistance, encodes a
phosphoglucosamine mutase."
RL J. Bacteriol. 179:5321-5325(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=COL;
RX MEDLINE=99265121; PubMed=10332717;
RA Wu S., de Lencastre H.;
RT "Mrp-a new auxiliary gene essential for optimal expression of
methicillin resistance in Staphylococcus aureus."
RL Microb. Drug Resist. 5:9-18(1999).
DR EMBL; Y09927; CAB55329.1; -;
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
SQ SEQUENCE 2478 AA; 263031 MW; 6B9859A02D023C74 CRC64;

Query Match 2.0%; Score 8; DB 2; Length 2478;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 97 AQPAAPVA 104
Db 161 AQPAAPVA 168

```
RESULT 65
Q9LCH2 PRELIMINARY; PRT; 2478 AA.
AC Q9LCH2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE FmtB.
GN FMTB.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COL;
RX MEDLINE=20348625; PubMed=10896508;
RA Komatsuzaawa H., Ohta K., Sugai M., Fujiwara T., Glanzmann P.,
RA Berger-Bachi B., Suginata H.;
RT "In551-mediated insertion inactivation of the fmbB gene encoding a
RT cell wall-associated protein abolishes methicillin resistance in
RT Staphylococcus aureus.";
RL J. Antimicrob. Chemother. 45:421-431(2000).
DR EMBL; AB025716; BAA93438.1; -
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF04650; YsIRK signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG anchor; 1.
DR TIGRFAMs; TIGR01168; YsIRK signal; 1.
SQ SEQUENCE 2478 AA; 262995 MW; 1C118EBE0DB03B34 CRC64;

Query Match 2.0%; Score 8; DB 2; Length 2478;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 AQAAPVA 104
Db 161 AQAAPVA 168

RESULT 66
Q99QR6 PRELIMINARY; PRT; 2481 AA.
AC Q99QR6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE FmtB protein.
GN FMTB (MRP) OR FMTB OR SAV2160 OR SA1964.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878; 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iino T., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003364; BAB58322.1; -
DR EMBL; AP003136; BAB43253.1; -
DR InterPro; IPR000890; Acetate_kin.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF04650; YsIRK signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
```

```
DR TIGRFAMs; TIGR01168; YsIRK signal; 1.
DR PROSITE; PS01075; ACETATE_KINASE_1; 1.
SQ SEQUENCE 2481 AA; 263765 MW; E1EAAB99B81665E8 CRC64;

Query Match 2.0%; Score 8; DB 16; Length 2481;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 AQAAPVA 104
Db 144 AQAAPVA 151

RESULT 67
Q8QL53 PRELIMINARY; PRT; 2593 AA.
AC Q8QL53;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Non-structural polyprotein.
OS Sleeping disease virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus.
OX NCBI_TaxID=78540;
RN [1]
RP SEQUENCE FROM N.A.
RA Bremon M.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Weston J.;
RT "Comparison of two aquatic alphaviruses, Salmon Pancreas Disease Virus
RT and Sleeping Disease Virus, using genome sequence analysis, monoclonal
RT reactivity and cross-infection.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ316246; CAC87660.1; -
DR InterPro; IPR002589; Alpp.
DR InterPro; IPR002877; FtsJ.
DR InterPro; IPR002620; Peptidase C9.
DR InterPro; IPR001788; RNA_dep_RNAPol2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR InterPro; IPR000606; Viral_helicase1.
DR InterPro; IPR002588; V_methyltransf.
DR Pfam; PF01661; Alpp; 1.
DR Pfam; PF01728; FtsJ; 1.
DR Pfam; PF01707; Peptidase C9; 1.
DR Pfam; PF00978; RNA_dep_RNAPol2; 1.
DR Pfam; PF01443; Viral_helicase1; 1.
DR Pfam; PF01660; Vmethyltransf; 1.
DR SMART; SM00506; Alpp; 1.
DR PROSITE; PS05057; RDRP_POSITIVE; 1.
DR PROSITE; PS05052; RDRP_VIRAL; 1.
KW Polyprotein.
FT CHAIN 1 563 NON STRUCTURAL PROTEIN P1.
FT CHAIN 564 1420 NON STRUCTURAL PROTEIN P2.
FT CHAIN 1421 1984 NON STRUCTURAL PROTEIN P3.
FT CHAIN 1985 2593 NON STRUCTURAL PROTEIN P4.
SQ SEQUENCE 2593 AA; 284851 MW; 2E2B4F651A45B8CP CRC64;

Query Match 2.0%; Score 8; DB 12; Length 2593;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 APVAPAA 108
Db 1834 APVAPAA 1841

RESULT 68
Q9UIC3
```



```

ID Q9U1C3 PRELIMINARY; PRT; 2861 AA.
AC Q9U1C3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Uncertain, possible lysosomal trafficking regulator or transport
DE protein.
GN L2581.05.
OS Leishmania major.
OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RA Zimmermann W., Wambutt R., Ivens A.C., Lawson D., Quail M.,
RA Rajandream M.A., Barrett B.G.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
DR EMBL; AL121851; CAB58291.1; -.
DR InterPro; IPR000409; Beige_BEACH.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF02138; Beach; 1.
DR ProDom; PD007848; Beige_BEACH; 1.
DR PROSITE; PS0197; BEACH; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
SQ SEQUENCE 2861 AA; 312361 MW; 2639D42C48AC843C CRC64;

Query Match 2.0%; Score 8; DB 5; Length 2861;
Best Local Similarity 100.0%; Pred.No. 3.9e-02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 ASSSGAST 86
Db 1547 ASSSGAST 1554

RESULT 69
Q95M44 PRELIMINARY; PRT; 48 AA.
ID Q95M44
AC Q95M44;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Insulin-like growth factor receptor type 2 (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Pfaffl M.W., Mircheva-Georgieva T., Penchev-Georgiev I., Hageleit M.,
RA Blum J.W.;
RT "Real-time RT-PCR quantification of insulin-like growth factor (IGF)-
RT 1, IGF-1 receptor, IGF-2, IGF-2 receptor, insulin receptor, growth
RT hormone receptor, IGF-binding proteins 1, 2 and 3 in the bovine
RT species.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ320234; CAC44343.1; -.
DR InterPro; IPR000479; CIMR.
DR Pfam; PF00878; CIMR; 1.
KW Receptor.
FT NON_TER 1 1
SQ NON_TER 48 48
SQ SEQUENCE 48 AA; 5436 MW; 4A27CC2A96E23533 CRC64;

```

```

Query Match 1.7%; Score 7; DB 6; Length 48;
Best Local Similarity 100.0%; Pred.No. 75;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 333 EGGPGGN 339
Db 38 EGGPGGN 44

RESULT 70
Q9K8K3 PRELIMINARY; PRT; 58 AA.
ID Q9K8K3
AC Q9K8K3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein BH3003.
GN BH3003.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001517; BAB06722.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 58 AA; 6931 MW; 0F04E8CDD8CD6C47 CRC64;

Query Match 1.7%; Score 7; DB 16; Length 58;
Best Local Similarity 100.0%; Pred.No. 90;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 357 EEREAIQ 363
Db 10 EEREAIQ 16

RESULT 71
P79662 PRELIMINARY; PRT; 74 AA.
ID P79662
AC P79662;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE MHC class I alpha-3 (Fragment).
OS Oncorhynchus tshawytscha (Chinook salmon) (King salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=74940;
RN [1]
RP SEQUENCE OF 3-28 FROM N.A.
RC STRAIN=A3-Onts-HN-11; TISSUE=Liver;
RX MEDLINE=98009322; PubMed=9348703;
RA Miller K.M., Withler R.E., Beacham T.D.;
RT "Molecular evolution at Mhc genes in two populations of chinook salmon
RT Oncorhynchus tshawytscha.";
RL Mol. Ecol. 6:937-954(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3-Onts-HN-11; TISSUE=Liver;
RA Miller K.M.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U80312; AAC78436.1; -.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_cl.

```

```

DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00407; IgC1; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 74
SQ SEQUENCE 74 AA; 8344 MW; A5A5B3481ESB30FE CRC64;

Query Match
Best Local Similarity 1.7%; Score 7; DB 7; Length 74;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 352 LTVTPPE 358
DB 62 LTVTPPE 68

RESULT 72
P79652 ID P79652 PRELIMINARY; PRT; 74 AA.
AC STRAIN=A3-Onts-H-10; TISSUE=Liver;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE MHC class I alpha-3 (fragment).
OS Oncorhynchus tshawytscha (Chinook salmon) (King salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=74940;
RN [1]
RP SEQUENCE OF 3-28 FROM N.A.
RC STRAIN=A3-Onts-H-10; TISSUE=Liver;
RX MEDLINE=98009322; PubMed=9348703;
RA Miller K.M., Withler R.E., Beacham T.D.;
RT "Molecular evolution at Mhc genes in two populations of chinook salmon
ONcorhynchus tshawytscha."
RL Mol. Ecol. 6:937-954(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3-Onts-H-10; TISSUE=Liver;
RA Miller K.M.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U80302; AAC78426.1; -.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00407; IgC1; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 74
SQ SEQUENCE 74 AA; 8320 MW; BIASB74CIESB30F9 CRC64;

Query Match
Best Local Similarity 1.7%; Score 7; DB 7; Length 74;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 352 LTVTPPE 358
DB 62 LTVTPPE 68

RESULT 73
Q8XTY0 ID Q8XTY0 PRELIMINARY; PRT; 74 AA.
AC Q8XTY0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein RSC3415.
GN RSC3415 OR RS01795.
OS Ralstonia solanacearum (Pseudomonas solanacearum).

```

```

OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Ralstoniaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Sigrier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum."
RL Nature 415:497-502(2002).
DR EMBL; AL646075; CAD16912.1; -.
KW Hypothetical protein: Complete proteome.
SQ SEQUENCE 74 AA; 7846 MW; 1C0A9ABF0ECF3758 CRC64;

Query Match
Best Local Similarity 1.7%; Score 7; DB 16; Length 74;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 ASSSGAS 85
DB 2 ASSSGAS 8

RESULT 74
Q8L656 ID Q8L656 PRELIMINARY; PRT; 75 AA.
AC Q8L656;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN OJ1203D03.2 OR OJ1781E12.2.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
RA Saki C., Henry D., Oates R., Simmons J.;
RT "Rice Genomic Sequence."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Wing R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Currie J.,
RA Collura K.;
RT "Rice Genomic Sequence."
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC099732; AAM27465.1; -.
DR EMBL; AC105927; AAO06954.1; -.
DR Gramene; Q8L656; -.
KW Hypothetical protein.
SQ SEQUENCE 75 AA; 8396 MW; F5D1D76BDB3E1228 CRC64;

Query Match
Best Local Similarity 1.7%; Score 7; DB 10; Length 75;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 AKASSG 83
DB 28 AKASSG 34

RESULT 75
Q9L824 ID Q9L824 PRELIMINARY; PRT; 83 AA.
AC Q9L824;

```

DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Hypothetical 9.6 kDa protein.
OS Enterococcus faecalis (Streptococcus faecalis).
OG Plasmid pIP834.
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BM4382; TRANSPOSON=Tn1549;
RA Garnier F., Taourit S., Glaser P., Courvalin P., Galimand M.;
RT "Characterization of transposon Tn1549 conferring VanB-type resistance
in Enterococcus sp.";
RL Microbiology 0:0-0(2000).
DR EMBL: AF192329; AAF72349.1; -.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 83 AA; 9602 MW; 323245C7D9CCEEF7 CRC64;

Query Match 1.7%; Score 7; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 131 QAAPAAT 137
Db 65 QAAPAAT 71
|||||

Search completed: December 17, 2003, 06:35:14
Job time : 46.8189 secs

Wed Dec 15 1955

83 7 1.7 65 22 ABG29760 Novel human diageno
84 7 1.7 66 22 AAU59913 Propionibacterium
85 7 1.7 66 22 ABP34520 Human ORF3493 prot
86 7 1.7 71 19 AAW72194 HSV-2 strain SB5 C
87 7 1.7 80 22 ABG49912 Human liver peptid
88 7 1.7 80 22 ABB28890 Peptide #2541 enco
89 7 1.7 80 22 ABB35070 Peptide #2576 enco
90 7 1.7 80 22 ABB20486 Protein #2485 enco
91 7 1.7 80 22 AAM55885 Human brain expres
92 7 1.7 80 22 AAM69258 Human bone marrow
93 7 1.7 80 22 AAM16081 Peptide #2515 enco
94 7 1.7 80 22 AAM28574 Peptide #2611 enco
95 7 1.7 80 22 AAM3808 Peptide #2490 enco
96 7 1.7 80 23 AAG37793 Human peptid enco
97 7 1.7 81 22 AAU59644 Propionibacterium
98 7 1.7 85 16 AAR71469 Premature thermoph
99 7 1.7 85 21 AAG30527 Arabidopsis thalia
100 7 1.7 89 22 AAU49184 Propionibacterium

ALIGNMENTS

RESULT 1

AA71458
ID AA71458 standard; Protein; 405 AA.
AC AA71458;

DT 04-OCT-2000 (first entry)
DE Maize Rad23 protein #1.

XX Rad23; maize; ATCC No: PTA-530; recombinant expression cassette; wheat;
KW transgenic plant; soybean; sunflower; sorghum; canola; modulator.
XX Zea mays.

OS WO200031268-A1.
PN 02-JUN-2000.

PD 12-OCT-1999; 99WO-US24129.
PF 23-NOV-1998; 98US-0109728.

PR (PION-) PIONEER HI-BRED INT INC.
PA Mahajan PB, Tagliani L;

PI WPI; 2000-400078/34.
DR N-PSDB; AAD01230.

XX Isolated nucleic acid encoding maize RAD23 protein is used to modulate
PT the levels of polypeptides in plant or in assays for identifying
PT compounds that bind to and/or increase/decrease enzymatic activity of
PT catalytically active polypeptides -
XX Claim 11b; Page 75-76; 82pp; English.

XX The present sequence is the maize Rad23 protein #1. It is isolated from
CC V5 root tissue of a Zea mays cell line B73, infected with corn root
CC worm. The cDNA sequence is deposited under the ATCC No: PTA-530. Maize
CC Rad23 DNA sequence operably linked to a promoter can be used to construct
CC a recombinant expression cassette. This expression cassette can be used
CC to generate a dicot or monocot transgenic plant e.g., maize, soybean,
CC sunflower, sorghum, canola, wheat, etc.. It can also be used to modulate
CC the levels of Rad23 polypeptide expression in a plant or in assays to
CC identify compounds, that bind to and/or modulate the enzymatic activity
CC of catalytically active polypeptides.

XX Sequence 405 AA;

Query Match 100.0%; Score 405; DB 21; Length 405;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 M K L N V K T L K G T N F I E A S P D A S V A D V K R I I E T T Q G S T Y R A D Q Q M L I Y Q G K I L K D E T T L E 60
DB 1 M K L N V K T L K G T N F I E A S P D A S V A D V K R I I E T T Q G S T Y R A D Q Q M L I Y Q G K I L K D E T T L E 60
QY 61 S N G V A E N S F L V I M L S K A S S G A S T A T T A K A P A T L A Q P A P A P A A S V A R T P T Q A P V A T 120
DB 61 S N G V A E N S F L V I M L S K A S S G A S T A T T A K A P A T L A Q P A P A P A A S V A R T P T Q A P V A T 120
QY 121 A E T A P P S V Q P Q A A P A A T V A A T D D A D V Y S Q A A S N I V F G N N L E O T I Q I L D M G G G T W E R D T V 180
DB 121 A E T A P P S V Q P Q A A P A A T V A A T D D A D V Y S Q A A S N I V F G N N L E O T I Q I L D M G G G T W E R D T V 180
QY 181 V R A L R A A Y N N P E R A I D Y L Y S G I P E N V E A Q P V A R A P A A G Q O T N Q A A S P A Q A P A L P V Q P S 240
DB 181 V R A L R A A Y N N P E R A I D Y L Y S G I P E N V E A Q P V A R A P A A G Q O T N Q A A S P A Q A P A L P V Q P S 240
QY 241 P A S A G P N A N P L N L P R O G V P S G S N P G V F G A S G A L D A L R Q L P Q F Q A L L Q L V Q A N P Q I L Q 300
DB 241 P A S A G P N A N P L N L P R O G V P S G S N P G V F G A S G A L D A L R Q L P Q F Q A L L Q L V Q A N P Q I L Q 300
QY 301 P M L Q E L G K N P Q I L R I Q E N Q A E F L R L V N E S P E G G P G G N I L G Q L A A A V P O T L T V T P E E R E 360
DB 301 P M L Q E L G K N P Q I L R I Q E N Q A E F L R L V N E S P E G G P G G N I L G Q L A A A V P O T L T V T P E E R E 360
QY 361 A I Q R L E G M G F N R E L V L E V F F A C N K D E L T A N Y L L D H G H E F D D Q Q Q 405
DB 361 A I Q R L E G M G F N R E L V L E V F F A C N K D E L T A N Y L L D H G H E F D D Q Q Q 405

RESULT 2

AA619576
ID AAG19976 standard; Protein; 307 AA.
AC AAG19976;

DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 21983.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.
OS Arabidopsis thaliana.

PN EP1033405-A2.
XX 06-SEP-2000.

PD 25-FEB-2000; 2000EP-0301439.
PF 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.
PR 18-APR-1999; 99US-0128714.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 08-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140931.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149802.
PR 23-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.

```
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 5.2%; Score 21; DB 21; Length 307;
Best Local Similarity 100.0%; Pred. No. 5e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 292 VQANPQILOPMLQELGKQNPQ 312
Db 198 VQANPQILOPMLQELGKQNPQ 218

RESULT 3
AAG45206
ID AAG45206 standard; Protein; 307 AA.
XX
AC AAG45206;
XX
DT 18-OCT-2000 DT (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 56725.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 21-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 25-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
```

PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160983.
PR 25-OCT-1999; 99US-0161404.

PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 5.2%; Score 21; DB 21; Length 307;
Best Local Similarity 100.0%; Pred. No. 5e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 292 VQANPQILQPMQLQELGKQNPQ 312
|||
DB 198 VQANPQILQPMQLQELGKQNPQ 218
|||

RESULT 4
AAG19975
ID AAG19975 standard; Protein; 335 AA.
XX
AC AAG19975;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 21982.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158365.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.

```
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      5.2%; Score 21; DB 21; Length 335;
Best Local Similarity 100.0%; Pred. No. 5.5e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 292 VQANPOILQPMLOELGKQNPQ 312
Db 226 VQANPOILQPMLOELGKQNPQ 246

RESULT 5
AAG45205
ID AAG45205 standard; Protein; 335 AA.
XX
AC AAG45205;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 56724.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000BP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 28-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 06-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139464.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
```

PR	09-AUG-1999;	99US-0147493.	
PR	09-AUG-1999;	99US-0147935.	
PR	10-AUG-1999;	99US-0148171.	
PR	11-AUG-1999;	99US-0148319.	
PR	12-AUG-1999;	99US-0148341.	
PR	13-AUG-1999;	99US-0148565.	
PR	13-AUG-1999;	99US-0148684.	
PR	16-AUG-1999;	99US-0149368.	
PR	17-AUG-1999;	99US-0149175.	
PR	18-AUG-1999;	99US-0149426.	
PR	20-AUG-1999;	99US-0149722.	
PR	20-AUG-1999;	99US-0149723.	
PR	20-AUG-1999;	99US-0149929.	
PR	23-AUG-1999;	99US-0149902.	
PR	23-AUG-1999;	99US-0149930.	
PR	25-AUG-1999;	99US-0150566.	
PR	26-AUG-1999;	99US-0150884.	
PR	27-AUG-1999;	99US-0151065.	
PR	27-AUG-1999;	99US-0151066.	
PR	27-AUG-1999;	99US-0151080.	
PR	30-AUG-1999;	99US-0151303.	
PR	31-AUG-1999;	99US-0151438.	
PR	01-SEP-1999;	99US-0151930.	
PR	07-SEP-1999;	99US-0152363.	
PR	10-SEP-1999;	99US-0153070.	
PR	13-SEP-1999;	99US-0153758.	
PR	15-SEP-1999;	99US-0154018.	
PR	16-SEP-1999;	99US-0154039.	
PR	20-SEP-1999;	99US-0154779.	
PR	22-SEP-1999;	99US-0155139.	
PR	23-SEP-1999;	99US-0155486.	
PR	24-SEP-1999;	99US-0155658.	
PR	28-SEP-1999;	99US-0156458.	
PR	29-SEP-1999;	99US-0156596.	
PR	04-OCT-1999;	99US-0157117.	
PR	05-OCT-1999;	99US-0157753.	
PR	06-OCT-1999;	99US-0157865.	
PR	07-OCT-1999;	99US-0158029.	
PR	08-OCT-1999;	99US-0158232.	
PR	12-OCT-1999;	99US-0158369.	
PR	13-OCT-1999;	99US-0159293.	
PR	13-OCT-1999;	99US-0159294.	
PR	13-OCT-1999;	99US-0159295.	
PR	14-OCT-1999;	99US-0159329.	
PR	14-OCT-1999;	99US-0159330.	
PR	14-OCT-1999;	99US-0159331.	
PR	14-OCT-1999;	99US-0159637.	
PR	14-OCT-1999;	99US-0159638.	
PR	18-OCT-1999;	99US-0159584.	
PR	21-OCT-1999;	99US-0160741.	
PR	21-OCT-1999;	99US-0160767.	
PR	21-OCT-1999;	99US-0160768.	
PR	21-OCT-1999;	99US-0160770.	
PR	21-OCT-1999;	99US-0160814.	
PR	21-OCT-1999;	99US-0160815.	
PR	22-OCT-1999;	99US-0160980.	
PR	22-OCT-1999;	99US-0160981.	
PR	22-OCT-1999;	99US-0160989.	
PR	25-OCT-1999;	99US-0161404.	
PR	25-OCT-1999;	99US-0161405.	
PR	25-OCT-1999;	99US-0161406.	
PR	26-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	
PR	26-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161320.	
PR	28-OCT-1999;	99US-0161992.	
PR	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	

Query Match 5.2%; Score 21; DB 21; Length 335;
Best Local Similarity 100.0%; Pred. No. 5.5e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	292	VOANFQILQPMQLQELGKNPQ	312
Db	226	VOANFQILQPMQLQELGKNPQ	246
RESULT 6			
AAG19974			
ID	AAG19974	standard; Protein; 378 AA.	
XX			
AC	AAG19974;		
XX			
DT	17-OCT-2000	(first entry)	
XX			
DE	Arabidopsis thaliana	protein fragment SEQ ID NO: 21981.	
XX			
KW	Protein identification; signal transduction pathway; metabolic pathway;		
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		
KW	termination sequence.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	EP1033405-A2.		
XX			
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000; 2000EP-0301439.		
XX			
PR	25-FEB-1999;	99US-0121825.	
PR	05-MAR-1999;	99US-0123180.	
PR	09-MAR-1999;	99US-0123548.	
PR	23-MAR-1999;	99US-0125788.	
PR	25-MAR-1999;	99US-0126264.	
PR	29-MAR-1999;	99US-0126785.	
PR	01-APR-1999;	99US-0127462.	
PR	06-APR-1999;	99US-0128234.	
PR	08-APR-1999;	99US-0128714.	
PR	16-APR-1999;	99US-0129845.	
PR	19-APR-1999;	99US-0130077.	
PR	21-APR-1999;	99US-0130449.	
PR	23-APR-1999;	99US-0130510.	
PR	23-APR-1999;	99US-0130891.	
PR	28-APR-1999;	99US-0131449.	
PR	30-APR-1999;	99US-0132048.	
PR	30-APR-1999;	99US-0132407.	
PR	04-MAY-1999;	99US-0132484.	
PR	05-MAY-1999;	99US-0132485.	
PR	06-MAY-1999;	99US-0132486.	
PR	07-MAY-1999;	99US-0132487.	
PR	11-MAY-1999;	99US-0132863.	
PR	11-MAY-1999;	99US-0134256.	
PR	14-MAY-1999;	99US-0134218.	
PR	14-MAY-1999;	99US-0134219.	
PR	14-MAY-1999;	99US-0134221.	
PR	14-MAY-1999;	99US-0134370.	
PR	19-MAY-1999;	99US-0134941.	
PR	20-MAY-1999;	99US-0135124.	
PR	21-MAY-1999;	99US-0135353.	
PR	24-MAY-1999;	99US-0135629.	
PR	25-MAY-1999;	99US-0136021.	
PR	27-MAY-1999;	99US-0136392.	
PR	28-MAY-1999;	99US-0136782.	
PR	01-JUN-1999;	99US-0137222.	
PR	03-JUN-1999;	99US-0137528.	
PR	04-JUN-1999;	99US-0137502.	
PR	07-JUN-1999;	99US-0137724.	
PR	08-JUN-1999;	99US-0138094.	
PR	10-JUN-1999;	99US-0138540.	
PR	10-JUN-1999;	99US-0138847.	
PR	14-JUN-1999;	99US-0139119.	
PR	16-JUN-1999;	99US-0139452.	
PR	16-JUN-1999;	99US-0139453.	
PR	17-JUN-1999;	99US-0139492.	

PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151085.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 5.2%; Score 21; DB 21; Length 378;
Best Local Similarity 100.0%; Pred. No. 6.2e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 292 VOANPQILQPMQLBGLGKNPQ 312

Db 269 VOANPQILQPMQLBGLGKNPQ 289

RESULT 7
AAG45204

ID AAG45204 standard; Protein; 378 AA.
XX AAG45204;
AC AAG45204;
XX AAG45204;
DT 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 56723.
DE Arabidopsis thaliana.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX EP1033405-A2.
PN EP1033405-A2.
XX EP1033405-A2.
PD 06-SEP-2000.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132408.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.

```
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      5.2%; Score 21; DB 21; Length 378;
Best Local Similarity 100.0%; Pred. No. 6.2e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 292 VOANPQILQPMQLQELGKQNPQ 312
Db 269 VOANPQILQPMQLQELGKQNPQ 289
|||||
|||||

RESULT 8
AAG4343
ID AAG44343 standard; Protein; 296 AA.
XX
AC
AAG44343;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55533.
```

```
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
```

```
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
```

```
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 23-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158234.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 4.4%; Score 18; DB 21; Length 296;
Best Local Similarity 100.0%; Pred. No. 3.2e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 295 NPQILQPLMLQELGKQNPQ 312
Db 188 NPQILQPLMLQELGKQNPQ 205
```

RESULT 9

AAG17379

ID AAG17379 standard; Protein; 299 AA.

XX

AC AAG17379;

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 18374.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0121180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136788.
PR 01-JUN-1999; 99US-0137228.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137503.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139499.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139765.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.


```
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155569.
PR 28-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0158029.
PR 07-OCT-1999; 99US-0158232.
PR 08-OCT-1999; 99US-0158369.
PR 12-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 4.4%; Score 18; DB 21; Length 299;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 295 NPQLQPMQLQELGKQNPQ 312
DB 191 NPQLQPMQLQELGKQNPQ 208

RESULT 10
AAG44342
ID AAG44342 standard; Protein; 345 AA.
XX AC AAG44342;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 55532.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
```

```
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0125785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
```

PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 17-JUL-1999; 99US-0144086.
PR 18-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 08-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.

PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.

Query Match 4.4%; Score 18; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 3.7e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 295 NPQILQPMQLQELGKQNPQ 312
|||
DB 237 NPQILQPMQLQELGKQNPQ 254

RESULT 11

AAG17378
ID AAG17378 standard; Protein; 348 AA.

XX AAG17378;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 18373.

XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; Genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 23-MAR-1999; 99US-0123548.

PR 25-MAR-1999; 99US-0125788.

PR 29-MAR-1999; 99US-0126264.

PR 01-APR-1999; 99US-0126785.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0128714.
PR 19-JUL-1999; 99US-0129845.
PR 19-JUL-1999; 99US-0130077.
PR 20-JUL-1999; 99US-0130449.
PR 20-JUL-1999; 99US-0130510.
PR 20-JUL-1999; 99US-0130891.
PR 21-JUL-1999; 99US-0131449.
PR 21-JUL-1999; 99US-0132048.
PR 21-JUL-1999; 99US-0132407.
PR 22-JUL-1999; 99US-0132484.
PR 22-JUL-1999; 99US-0132485.
PR 22-JUL-1999; 99US-0132486.
PR 22-JUL-1999; 99US-0132487.
PR 23-JUL-1999; 99US-0132863.
PR 23-JUL-1999; 99US-0134256.
PR 23-JUL-1999; 99US-0134218.
PR 23-JUL-1999; 99US-0134219.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.

```
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 4.4%; Score 18; DB 21; Length 348;
Best Local Similarity 100.0%; Pred.No. 3.8e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 295 NPQILQPMLELQKGNPQ 312
   |||||
Db 240 NPQILQPMLELQKGNPQ 257

RESULT 12
AAG44341
ID AAG44341 standard; Protein; 368 AA.
XX
AC AAG44341;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55531.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 08-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
```

```
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
```

PR	21-OCT-1999;	99US-0160741;
PR	21-OCT-1999;	99US-0160761;
PR	21-OCT-1999;	99US-0160768;
PR	21-OCT-1999;	99US-0160770;
PR	21-OCT-1999;	99US-0160815;
PR	21-OCT-1999;	99US-0160815;
PR	21-OCT-1999;	99US-0160880;
PR	22-OCT-1999;	99US-0160981;
PR	22-OCT-1999;	99US-0160989;
PR	22-OCT-1999;	99US-0161404;
PR	25-OCT-1999;	99US-0161405;
PR	25-OCT-1999;	99US-0161406;
PR	25-OCT-1999;	99US-0161359;
PR	26-OCT-1999;	99US-0161360;
PR	26-OCT-1999;	99US-0161361;
PR	28-OCT-1999;	99US-0161920;
PR	28-OCT-1999;	99US-0161992;
PR	28-OCT-1999;	99US-0161993;
PR	29-OCT-1999;	99US-0162142;
PR	29-OCT-1999;	99US-0162142;

Query Match 4.4%;
Best Local Similarity 100.0%;
Matches 18; Conservative

Qy	295	NPQILQPMQLGKGNPQ
Db	260	NPQILQPMQLGKGNPQ

RESULT 13
AAG17377

ID	AAG17377	standard; Protein;
XX	AC	
XX	AAG17377;	
XX		
DT	17-OCT-2000	(first entry)
XX		
DE	Arabidopsis thaliana protein	
XX		
KW	Protein identification; signa	
KW	hybridization assay; genet	
KW	termination sequence.	
XX		
OS	Arabidopsis thaliana.	
XX		
PN	EP1033405-A2.	
XX		
FD	06-SEP-2000.	
XX		
PF	25-FEB-2000;	2000EP-0301439.
XX		
XX	25-FEB-1999;	99US-0121825;
PR	05-MAR-1999;	99US-0123180;
PR	03-MAR-1999;	99US-0123548;
PR	23-MAR-1999;	99US-0125788;
PR	25-MAR-1999;	99US-0126264;
PR	29-MAR-1999;	99US-0126785;
PR	01-APR-1999;	99US-0127462;
PR	06-APR-1999;	99US-0128234;
PR	08-APR-1999;	99US-0128714;
PR	16-APR-1999;	99US-0129845;
PR	19-APR-1999;	99US-0130077;
PR	21-APR-1999;	99US-0130449;
PR	23-APR-1999;	99US-0130510;
PR	23-APR-1999;	99US-0130891;
PR	28-APR-1999;	99US-0131449;
PR	30-APR-1999;	99US-0132048;
PR	04-MAY-1999;	99US-0132407;
PR	04-MAY-1999;	99US-0132484;
PR	05-MAY-1999;	99US-0132485;
PR	06-MAY-1999;	99US-0132486;
PR	06-MAY-1999;	99US-0132487;
PR	07-MAY-1999;	99US-0132863;

PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 25-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158232.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.

```
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161362.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0162142.
PR 28-OCT-1999; 99US-0162143.
PR 28-OCT-1999; 99US-0162144.

Query Match 4.4% Score 18; DB 21; Length 371;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 295 NPQILQPMQLQELGKQNFQ 312
Db 263 NPQILQPMQLQELGKQNFQ 280

RESULT 14
AAG36531
ID AAG36531 standard; Protein; 186 AA.
XX
AC AAG36531;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 44781.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127482.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.

20-MAY-1999; 99US-0135124.
21-MAY-1999; 99US-0135353.
24-MAY-1999; 99US-0135629.
25-MAY-1999; 99US-0136021.
27-MAY-1999; 99US-0136392.
28-MAY-1999; 99US-0136782.
01-JUN-1999; 99US-0137222.
03-JUN-1999; 99US-0137528.
04-JUN-1999; 99US-0137502.
07-JUN-1999; 99US-0137724.
08-JUN-1999; 99US-0138094.
10-JUN-1999; 99US-0138540.
10-JUN-1999; 99US-0138847.
14-JUN-1999; 99US-0139119.
16-JUN-1999; 99US-0139452.
16-JUN-1999; 99US-0139453.
17-JUN-1999; 99US-0139492.
18-JUN-1999; 99US-0139454.
18-JUN-1999; 99US-0139455.
18-JUN-1999; 99US-0139456.
18-JUN-1999; 99US-0139457.
18-JUN-1999; 99US-0139458.
18-JUN-1999; 99US-0139459.
18-JUN-1999; 99US-0139460.
18-JUN-1999; 99US-0139461.
18-JUN-1999; 99US-0139462.
18-JUN-1999; 99US-0139463.
18-JUN-1999; 99US-0139750.
18-JUN-1999; 99US-0139763.
21-JUN-1999; 99US-0139817.
22-JUN-1999; 99US-0139899.
23-JUN-1999; 99US-0140353.
24-JUN-1999; 99US-0140354.
24-JUN-1999; 99US-0140695.
28-JUN-1999; 99US-0140823.
29-JUN-1999; 99US-0140991.
30-JUN-1999; 99US-0141287.
01-JUL-1999; 99US-0141842.
01-JUL-1999; 99US-0142154.
02-JUL-1999; 99US-0142055.
06-JUL-1999; 99US-0142390.
08-JUL-1999; 99US-0142803.
09-JUL-1999; 99US-0142920.
12-JUL-1999; 99US-0142977.
13-JUL-1999; 99US-0143542.
14-JUL-1999; 99US-0143624.
15-JUL-1999; 99US-0144005.
16-JUL-1999; 99US-0144085.
16-JUL-1999; 99US-0144086.
19-JUL-1999; 99US-0144325.
19-JUL-1999; 99US-0144331.
19-JUL-1999; 99US-0144332.
19-JUL-1999; 99US-0144333.
19-JUL-1999; 99US-0144334.
19-JUL-1999; 99US-0144335.
20-JUL-1999; 99US-0144352.
20-JUL-1999; 99US-0144632.
20-JUL-1999; 99US-0144884.
21-JUL-1999; 99US-0144814.
21-JUL-1999; 99US-0145086.
21-JUL-1999; 99US-0145088.
22-JUL-1999; 99US-0145085.
22-JUL-1999; 99US-0145087.
22-JUL-1999; 99US-0145089.
22-JUL-1999; 99US-0145192.
23-JUL-1999; 99US-0145145.
23-JUL-1999; 99US-0145218.
23-JUL-1999; 99US-0145224.
26-JUL-1999; 99US-0145276.
27-JUL-1999; 99US-0145513.
27-JUL-1999; 99US-0145918.
27-JUL-1999; 99US-0145919.
28-JUL-1999; 99US-0145951.
```

```
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149502.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 21-SEP-1999; 99US-0155486.
PR 21-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 3.2%; Score 13; DB 21; Length 186;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 245 GFNANPLNLFPOG 257
DB 154 GFNANPLNLFPOG 166
|||||

RESULT 15
AAG36530
ID AAG36530 standard; Protein; 214 AA.
XX
AC AAG36530;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 44780.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
```


PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 08-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145921.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

3.2%; Score 13; DB 21; Length 214;

Best Local Similarity 100.0%; Pred. No. 0.0012; Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 245 GPNANPLNLFPOG 257
Db 182 GPNANPLNLFPOG 194

RESULT 16

AAG36529
ID AAG36529 standard; Protein; 257 AA.

XX
AC AAG36529;
DT 18-OCT-2000 (first entry)
XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 44779.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-012180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145921.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147152.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149302.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150366.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 3.2%; Score 13; DB 21; Length 257;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 245 GPNANPLNLPQ 257
Db 225 GPNANPLNLPQ 237

RESULT 17
AAY71459

ID AAY71459 standard; Protein; 368 AA.

XX AC AAY71459;

XX DT 04-OCT-2000 (first entry)

XX DE Maize Rad23 protein #2.

KW Rad23; maize; ATCC No: PTA-531; recombinant expression cassette; wheat;
KW transgenic plant; soybean; sunflower; sorghum; canola; modulator.

OS Zea mays.

PN WO200031268-A1.

XX PD 02-JUN-2000.

XX PF 12-OCT-1999; 99WO-US24129.

XX PR 23-NOV-1998; 98US-0109728.

XX PA (PION-) PIONEER HI-BRED INT INC.

XX PI Mahajan PB, Tagliani L;

DR WPI; 2000-400078/34.

DR N-PSDB; AAD01231.

PT Isolated nucleic acid encoding maize RAD23 protein is used to modulate
the levels of polypeptides in plant or in assays for identifying
PT compounds that bind to and/or increase/decrease enzymatic activity of
PT catalytically active polypeptides -

XX Claim 11b; Page 78-79; 82pp; English.

XX The present sequence is the maize Rad23 protein #2. It is isolated from
a Zea mays cell line, B73 callus tissue regenerated five days after
transfer of the callus from medium containing auxin to a medium devoid
of exogenous auxin. The cDNA is deposited under the ATCC No: PTA-531.
CC Maize Rad23 DNA sequence operably linked to a promoter, can be used to
construct a recombinant expression cassette. This expression cassette
can be used to generate a dicot or monocot transgenic plant e.g., maize,
soybean, sunflower, sorghum, canola, wheat, etc. It can also be used to
modulate the levels of Rad23 polypeptide expression in a plant or in
assays to identify compounds, that bind to and/or modulate the enzymatic
activity of catalytically active polypeptides.

XX SQ Sequence 368 AA;

Query Match 3.2%; Score 13; DB 21; Length 368;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 182 RALRAAYNNPERA 194

Db 167 RALRAAYNNPERA 179

RESULT 18

ABP34618

ID ABP34618 standard; Protein; 117 AA.

XX AC ABP34618;

XX DT 08-JUL-2002 (first entry)

XX DE Human ORF3591 protein, SEQ ID NO:7182.

KW Human; ORF; open reading frame; ORF; drug screening; diagnosis;
KW disease monitoring; cytokine; cell proliferation; cell differentiation;

immune modulation; haematopoiesis regulation; tissue growth; angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic; thrombolytic; tumour inhibition; bodily characteristic; fertility; behaviour; cancer; proliferative disorder; neurological disorder; cardiovascular disease; immune system disorder; organ transplantation; tissue growth disorder; tissue regeneration disorder; diabetes mellitus; hypothyroidism; cholesterol ester storage disease; infection; vulnery; vasotropic; antipruritic; antidiabetic; cytostatic; nootropic; neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic; cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator; dermatological; analgesic; virucide; antibacterial; fungicide.

Homo sapiens.

WO200190366-A2.

29-NOV-2001.

24-MAY-2001; 2001WO-US17076.

24-MAY-2000; 2000US-206690P.

(CURA-) CURAGEN CORP.

Leach MD, Shinkets RA;

WPI; 2002-106200/14.

N-PSDB; ABN78644.

Novel human polypeptides and polynucleotides useful for diagnosing, preventing and treating cardiovascular disease, neurodegenerative, hyperproliferative disorders and disorders related to organ transplantation -

Claim 10; Page 2042; 2508pp; English.

Sequences ABP31028-ABP3561 represent 4534 novel human proteins designated ORF (open reading frame) 1-4534, and sequences ABN75054-ABN79587 represent cDNAs encoding them. The invention also encompasses polypeptides at least 80% identical to the ORF1-ORF4534 (collectively referred to as ORFX) proteins, polynucleotides at least 85% identical to the ORFX nucleic acid sequences, vectors and host cells comprising ORFX polynucleotides, the recombinant production of ORFX proteins, antibodies specific for ORFX proteins, methods of detecting ORFX polynucleotides and polypeptides, methods of screening for modulators of ORFX expression or activity, and methods of screening individuals for a predisposition to an ORFX-associated disorder. The ORFX proteins of the invention have a wide range of biological activities, such as cytokines, cell proliferation, cell differentiation, immune modulation, haematopoiesis regulation, tissue growth, angiogenesis, activin or inhibin activity, chemotactic/chemokinetic activity, haemostatic activity, thrombolytic activity, receptor/ligand, antiinflammatory activity, tumour inhibition activity, and antiinfective activity, and may also be involved in the determination of bodily characteristics, fertility and behaviour. ORFX proteins, nucleic acids and antibodies may be used in the treatment of cancers, other proliferative disorders such as psoriasis and benign tumours, neurological disorders such as epilepsy and Alzheimer's disease, cardiovascular diseases, immune system disorders, disorders related to organ transplantation, disorders of tissue growth and regeneration, diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester storage disease, and infectious diseases caused by viral, bacterial, fungal and other pathogens. ORFX nucleic acids may also be used as a source of primers and probes, in the detection of ORFX genomic sequences or transcripts, in the identification and cloning of homologous sequences, in genetic diagnosis, and in forensic biology. The ORFX nucleic acids may additionally be used to produce transgenic animals which may be useful for studying the function and/or activity of ORFX protein, and in drug screening. The ORFX proteins may also be used as immunogens to generate specific antibodies, which are useful in the diagnosis, treatment and monitoring of ORFX-associated diseases.

Sequence 117 AA;

Query Match 2.5%; Score 10; DB 23; Length 117;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 KQNPOILRLI 317
DB 51 KQNPOILRLI 60
|||||

RESULT 19

ABB69580
ID ABB69580 standard; Protein; 487 AA.

XX ABB69580;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 35532.

XX Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

XX 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

XX N-PSDB; ABL13683.

PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

XX Disclosure; SEQ ID NO 35532; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 487 AA;

Query Match 2.2%; Score 9; DB 22; Length 487;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 PAAPVAPAA 107

DB 158 PAAPVAPAA 166
|||||

RESULT 20

AAG81195
ID AAG81195 standard; Protein; 3070 AA.

XX

AC AAG81195;
XX
DT 04-SEP-2001 (first entry)
XX
DE Mycobacterium tuberculosis potential drug target protein SEQ ID 246.
XX
KW Drug target; growth; organism viability; characterisation.
XX
OS Mycobacterium tuberculosis.
XX
PN WO200135317-A1.
XX
PD 17-MAY-2001.
XX
PF 13-NOV-2000; 2000WO-US311152.
XX
PR 12-NOV-1999; 99US-0165086.
XX
PR 12-NOV-1999; 99US-0165124.
XX
PR 01-FEB-2000; 2000US-0179531.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Eisenberg D, Rotstein SH, Marcotte EM;
XX
DR N-PSDB; AAH52046.
XX
XX
XX Identifying nucleotide or polypeptide sequence for use as drug target,
PT involves providing algorithm that analyzes a functional relationship
PT between nucleotide or polypeptide sequences, and comparing the
PT sequences -
XX
XX Disclosure; Page 178-180; 207pp; English.
XX
XX This invention relates to a method for identifying a nucleotide or
CC polypeptide sequence that may be a drug target, or essential for growth
CC or viability of an organism. Polynucleotide sequences AAH51947 - AAH52092
CC represent DNA encoding proteins AAG81096 - AAG81241. Mycobacterium
CC tuberculosis proteins which are potential drug targets. The DNA and
CC protein sequences are used to illustrate the method of the invention. The
CC method involves providing an unknown nucleotide or polypeptide sequences,
CC and comparing it to a number of sequences along with at least one
CC algorithm capable of analysing a functional relationship between
CC nucleotide and polypeptide sequences. The method is useful for
CC characterising the function of nucleic acids and polypeptides that may be
CC useful as a target for a drug or essential for the growth or viability of
CC an organism.
XX
SQ Sequence 3070 AA;
Query Match 2.2%; Score 9; DB 22; Length 3070;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 100 AAPVAPPAAS 108
DB 1755 AAPVAPPAAS 1763
|||||
RESULT 21
AAR74046
ID AAR74046 standard; Peptide; 23 AA.
XX
AC AAR74046;
XX
XX 25-MAR-2003 (updated)
DT 01-DEC-1995 (first entry)
XX
XX Rice carbonic-anhydrase.
DE
XX Rice; carbonic-anhydrase; carbon-fixation; monocotyledon; cereal;
KW crop improvement.
XX

OS Oryza sativa.
XX
XX Key Location/Qualifiers
FT Misc-difference 17 /note= "amino acid at position 17 is not identified
FT in the specification"
XX
XX WO9511979-A1.
XX
XX 04-MAY-1995.
XX
XX 27-OCT-1994; 94WO-JP01814.
XX
XX 29-OCT-1993; 93JP-0294278.
XX
XX (NISR) JAPAN TOBACCO INC.
XX
XX Burnell NJ, Suzuki S;
PI WPI; 1995-178871/23.
XX
XX Cloned DNA coding for monocotyledon carbonic anhydrase - is used
PT to improve carbon fixation in monocotyledonous plants.
XX
XX Disclosure; Page 17; 42pp; Japanese.
XX
XX The cDNA sequence given in AAQ92067 was obt'd. by screening and
CC cloning a cDNA library generated from rice RNA, and encodes
CC carbonic-anhydrase (AAR74043). A peptide corresp. to amino acids
CC 63-85 of the encoded enzyme is given in AAR74046.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 23 AA;
Query Match 2.0%; Score 8; DB 16; Length 23;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 100 AAPVAPPA 107
DB 1 AAPVAPPA 8
|||||
RESULT 22
ABG12648
ID ABG12648 standard; Protein; 71 AA.
XX
XX ABG12648;
AC
XX 18-FEB-2002 (first entry)
DT
XX Novel human diagnostic protein #12639.
DE
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
OS
XX WO200175067-A2.
PN
XX 11-OCT-2001.
PD
XX 30-MAR-2001; 2001WO-US08631.
PF
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Drmanac RT, Liu C, Tang YT;
PI WPI; 2001-639362/73.
XX
XX N-PSDB; AAS76835.
DR

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 20; SEQ ID No 43007; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 71 AA;

Query Match 2.0%; Score 8; DB 22; Length 71;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 AAPVAPAA 107
Db 35 AAPVAPAA 42
|||||

RESULT 23
AAR99709
ID AAR99709 standard; Protein; 78 AA.
XX
AC AAR99709;
XX
DT 25-MAR-2003 (updated)
DT 26-SEP-1996 (first entry)
XX
DE pKTH1798-encoded protein.
XX
KW Promoter; signal sequence; protein secretion; vector;
KW Gram-positive bacterium; *Saccharichia coli*; *Lactobacillus*;
KW *Bacillus subtilis*; pKTH1798; beta-lactamase.
XX
OS *Lactococcus lactis* subsp. *lactis*.
XX
FH Key Location/Qualifiers
FT Region 77..78
FT /note= "N-terminus of TEM-beta-lactamase"
XX
FN US5529908-A.
XX
PD 25-JUN-1996.
XX
PF 10-FEB-1993; 93US-0015582.
XX
PR 10-JUL-1989; 89US-0377450.
PR 10-FEB-1993; 93US-0015582.
XX
PA (VALI-) VALIO LTD.
XX

PI Koivula T, Palva I, Sibakov M, Von Wright A;
XX WPI; 1996-308737/31.
DR N-PSDB; AAT31878.
XX
PT *Lactococcus* promoters and signal sequences - and related vectors and
PT transformed cells, provide efficient expression of proteins in Gram
PT positive bacteria
XX
XX Example 5; Fig 15; 20pp; English.
XX
CC *Lactococcus lactis* subsp. *lactis*-derived plasmid pKTH1798 (AAT31878)
CC codes for a protein (AAR99709) that includes the N-terminal region of
CC the TEM-beta-lactamase gene. The correct reading frame of pKTH1798
CC was detd. by matching the 3 reading frames with the known reading
CC frame of beta-lactamase. pKTH1798 promoter and signal sequences may
CC be incorporated into plasmids and used to achieve enhanced
CC heterologous gene expression and protein secretion in *E. coli* and
CC Gram-positive bacteria, esp. *Bacillus subtilis*, *Lactococcus* and
CC *Lactobacillus* spp.
CC (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ Sequence 78 AA;

Query Match 2.0%; Score 8; DB 17; Length 78;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 75 SKAKASSS 82
Db 33 SKAKASSS 40
|||||

RESULT 24
AAR22951
ID AAR22951 standard; Protein; 117 AA.
XX
AC AAR22951;
XX
DT 25-MAR-2003 (updated)
DT 15-OCT-1992 (first entry)
XX
DE Promoter/secretion promoting signal sequence from pKTH1797.
XX
KW Probe-vector; secretion; signal; promoter; plasmid; expression;
KW heterologous; homologous; *E. coli*; Gram-positive bacteria;
KW TEM; beta-lactamase; TEM-bla.
XX
OS *Lactobacillus lactis* subsp. *lactis*.
XX
FH Key Location/Qualifiers
FT Region 116..117
FT /note= "N-terminus of TEM-bla"
XX
FN WO9204451-A.
XX
PD 19-MAR-1992.
XX
PF 30-AUG-1990; 90WO-FI00204.
XX
PR 30-AUG-1990; 90WO-FI00204.
XX
XX (GENE-) GENESIT OY.
PA (VALI-) VALIO FINNISH COOPERATIVE DAIRIES ASSOC.
XX
PI Palva I, Sibakov M, Koivula T, Vonwright A, Sibakov MY;
XX WPI; 1992-131815/16.
DR N-PSDB; AAQ23883.
XX
XX Promoter probe vectors for protein expression - replicable in *E.*
XX *coli*, *B. subtilis*, *Lactococci* and *Lactobacillus*
XX

PS Disclosure; Fig 14; 70pp; English.

XX The three reading frames of the sequence represented in AAQ23883 were
CC matched with the known reading frame of beta-lactamase, in order to
CC determine the correct reading frame. The deduced amino acid
CC sequence of one reading frame is given here, and is an exact
CC reproduction of the sequence given in the specification, i.e.
CC including the seven X's which are encoded by stop codons.
CC Promoter probe-vectors pKTH1734, pKTH1736 and pKTH1750 were used to
CC clone and sequence previously unknown and undescribed L. lactis subsp.
CC lactis promoter and promoter/secretion signal promoting nucleotide
CC sequences. These sequences are contained in pKTH1816 (AAQ23878),
CC pKTH1817 (AAQ23879), pKTH1820 (AAQ23880), pKTH1874 (AAQ23881), pKTH1789
CC (AAQ23882), pKTH1797 (AAQ23883), pKTH1798 (AAQ23884), pKTH1799
CC (AAQ23885), pKTH1801 (AAQ23886), pKTH1821 (AAQ23887). The sequences may
CC be beneficially incorporated into plasmids, by means of which it is
CC possible to achieve enhanced heterologous (and homologous) protein
CC expression in E. coli and, esp., in Gram-positive bacteria.
CC (Updated on 25-MAR-2003 to correct PF field.)
CC (Updated on 25-MAR-2003 to correct PR field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
CC (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 117 AA;

Query Match 2.0%; Score 8; DB 13; Length 117;

Best Local Similarity 100.0%; Pred. No. 33;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 SKAKASSS 82

DB 72 SKAKASSS 79

RESULT 25

AA47659

ID AAR47659 standard; Protein; 117 AA.

XX

AC AAR47659;

XX

DT 25-MAR-2003 (updated)

DT 07-MAR-1994 (first entry)

XX

DE Promoter-secretion signal peptide from *Lactococcus lactis*.

XX Promoter; secretion; heterologous proteins; homologous proteins;

KW expression cassette; *Lactococcus lactis*.

XX

OS *Lactococcus lactis*.

XX

PN US5242821-A.

XX

PD 07-SEP-1993.

XX

PF 10-JUL-1989; 89US-0377450.

XX

PR 10-JUL-1989; 89US-0377450.

XX

PA (VALI-) VALIO.

XX

PI Palva I, Sibakov M, Koivula T, Von Wright A;

XX

DR WPI; 1993-295228/37.

XX

DR N-PSDB; AAQ46208.

XX

XX New promoters and secretion promoting signals - isolated from

PT *Lactococcus lactis*, used for expression of proteins in

PT Gram-positive bacteria

XX

PS Disclosure; Claim 4; Figure 15; 32pp; English.

XX

CC The *Lactococcus lactis* promoter-secretion signal coding sequence is

CC removed from the plasmid pKTH1798 and is operably linked to a

CC *Lactococcus lactis* promoter sequence. By using L. lactis promoters
CC and promoter/secretion signals, the production of heterologous and
CC homologous proteins in E.coli and gram-positive hosts can be greatly
CC improved.
CC (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 117 AA;

Query Match 2.0%; Score 8; DB 14; Length 117;

Best Local Similarity 100.0%; Pred. No. 33;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 SKAKASSS 82

DB 72 SKAKASSS 79

RESULT 26

AA26036

ID AAY26036 standard; Protein; 186 AA.

XX

AC AAY26036;

XX

DT 22-OCT-1999 (first entry)

XX

DE Secreted protein if87_1.

XX

KW Secreted protein; cytokine; cell proliferation; immune stimulation;

KW vaccine; immune suppression; haematopoiesis; tissue growth; activin;

KW inhibitor; chemotaxis; chemokinesis; haemostasis; thrombolytic;

KW receptor; ligand; anti-inflammatory; cadherin; tumour; gene therapy.

XX

OS *Homo sapiens*.

XX

PN WO9928335-A1.

XX

PD 10-JUN-1999.

XX

PF 02-DEC-1998; 98WO-US25512.

XX

PR 30-NOV-1998; 98US-0203106.

XX

PR 04-DEC-1997; 97US-0067454.

XX

PA (GENY) GENETICS INST INC.

XX

PI Agostino MJ, Clark HF, Collins-Racie LA, Evans C;

PI Fecthel K, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

XX

PI Steininger RJ, Treacy M, Wong GG;

XX

DR WPI; 1999-385352/32.

XX

DR N-PSDB; AAX80671.

XX

PT New polynucleotides encoding secreted human proteins

XX

PS Claim 25; Page 111; 124pp; English.

XX

CC The present sequence is a secreted protein if87_1, encoded by known clone

CC if87_1 (deposited as ATCC 98600) isolated from human adult uterus CDNA

CC library. Recombinant secreted proteins can be produced by

CC transforming host cells and culturing them under suitable conditions.

CC The polynucleotide and protein are predicted to have biological

CC activities which would make them suitable for treating, preventing or

CC ameliorating medical conditions in humans and animals. Some predicted

CC biological activities include cytokine and cell proliferation/

CC differentiation activity, immune stimulating (e.g. as vaccines) or

CC suppressing activity, haematopoiesis regulating activity, tissue growth

CC activity, activin/inhibin activity, chemotactic/chemokinetic activity,

CC hemostatic and thrombolytic activity, receptor/ligand activity, anti-

CC inflammatory activity, cadherin/tumour invasion suppressor activity, and

CC tumour inhibition activity. The polynucleotide encoding secreted

CC protein can be used for gene therapy.

XX Sequence 186 AA;

Query Match 2.0%; Score 8; DB 20; Length 186;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 SSGASTAT 88
 Db 39 SSGASTAT 46
 |||||

RESULT 27
 AAU39003
 ID AAU39003 standard; Protein; 186 AA.
 XX AC AAU39003;
 XX DT 16-JAN-2002 (first entry)
 XX DE Human secreted protein if87_1.
 XX KW Human; secreted protein; antiinflammatory; immunosuppressive;
 KW nontropic; neuroprotective; antiarthritic; antimicrobial; vulnery;
 KW cytosolic; antidiabetic; viricide; antifertility; anticonvulsant;
 KW vasotropic; antiparkinsonian; immunostimulant; dermatological;
 KW antirheumatic; antitumor; antitumor; osteopathic; tranquiliser;
 KW cerebroprotective; cytokine; cell proliferation; cell differentiation;
 KW immune deficiency; severe combined immunodeficiency; SCID; tumour;
 KW autoimmune disorder; multiple sclerosis; rheumatoid arthritis;
 KW graft-versus-host disease; myeloid deficiency; wound healing; ulcer;
 KW periodontal disease; osteoporosis; osteoarthritis; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; infection; cardiac disease;
 KW stroke; sepsis; inflammatory bowel disease; contraceptive; immunogen;
 KW food supplement; vaccine.
 XX OS Homo sapiens.
 XX FN WO200175068-A2.
 XX PD 11-OCT-2001.
 XX PF 22-MAR-2001; 2001WO-US09369.
 XX PR 30-MAR-2000; 2000US-0539330.
 XX PR 04-DEC-2000; 2000US-0729674.
 XX (GEMY) GENETICS INST INC.
 XX PA Jacobs K, McCoy JM, Lavallie E, Collins-racie LA, Evans C;
 PI Treacy M, Agostino MJ, Steininger RJ, Spaulding V, Wong GG;
 PI Clark H, Fechtel K, Merberg D;
 XX WPI; 2001-639363/73.
 XX DR N-PSDB; AAS59221.
 XX Secreted human proteins, useful as vaccine for treating various
 PT diseases such as autoimmune disorders (e.g. multiple sclerosis), and
 PT nervous system disorders (e.g. stroke) -
 XX Disclosure; Page 476; 619pp; English.
 XX The invention relates to novel human secreted proteins, the nucleic
 CC acids encoding them. The protein may exhibit cytokine, cell proliferation
 CC or cell differentiation activity or may induce production of other
 CC cytokines in certain cell populations and may exhibit immune stimulating
 CC or immune suppressing activity, which is useful for the treatment of
 CC various immune deficiencies and disorders e.g. severe combined
 CC immunodeficiency (SCID), autoimmune disorders e.g. multiple sclerosis,
 CC systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
 CC inflammation. The proteins are also useful in the treatment of diseases
 CC and disorders including tissue, skin and organ transplantation and in
 CC graft-versus-host diseases (GVHD), in the induction of tumour immunity,
 CC myeloid or lymphoid cell deficiencies, wound healing and tissue repair,
 CC in the treatment of burns, incisions and ulcers; as well as in treatment

CC of periodontal disease, osteoporosis or osteoarthritis, mediated by
 CC inflammatory processes, diseases of the peripheral nervous system,
 CC Alzheimer's, Parkinson's disease, Huntington's disease,
 CC amyotrophic lateral sclerosis, and Shy-Drager syndrome, infections,
 CC infarction of cardiac and central nervous system vessel e.g. stroke,
 CC sepsis, inflammatory bowel disease, ulcers, bone regeneration. The
 CC protein, having activin- or inhibin-related activities is useful as a
 CC contraceptive based on the ability of inhibin to decrease fertility in
 CC female mammals and decrease spermatogenesis in male mammals. The
 CC proteins and nucleic acids are also useful as food supplements. The
 CC present sequence represents a secreted protein of the invention.
 XX SQ Sequence 186 AA;
 Query Match 2.0%; Score 8; DB 22; Length 186;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 SSGASTAT 88
 Db 39 SSGASTAT 46
 |||||

RESULT 28
 ABB55712
 ID ABB55712 standard; Protein; 186 AA.
 XX AC ABB55712;
 XX DT 14-FEB-2002 (first entry)
 XX DE Human polypeptide SEQ ID NO 30.
 XX KW Human; clone bd306-7; clone yb8-1; ATCC number 98599; gene therapy;
 KW immune disorder; bacterial infection; fungal infection; cancer; tumour;
 KW autoimmune disorder; systemic lupus erythematosus; wound; ulcer; inhibin;
 KW osteoporosis; osteoarthritis; nervous system disorder; neuropathy;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; activin;
 KW haemophilia; cardiac infarction; stroke; sepsis; arthritis; vulnery;
 KW ischaemia-reperfusion injury; inflammatory bowel disease; chemotactic;
 KW Crohn's disease; cytostatic; anti-inflammatory; immunomodulator;
 KW neuroprotective; haemostatic; thrombolytic; anti-inflammatory.
 XX OS Homo sapiens.
 XX FN US2001039335-A1.
 XX PD 08-NOV-2001.
 XX PF 04-DEC-2000; 2000US-0729674.
 XX PR 10-APR-1997; 97US-126425P.
 XX PR 04-DEC-1997; 97US-067454P.
 XX PR 20-DEC-1997; 97US-068379P.
 XX PR 02-JAN-1998; 98US-070346P.
 XX PR 07-JAN-1998; 98US-070643P.
 XX PR 08-JAN-1998; 98US-070755P.
 XX PR 13-JAN-1998; 98US-071304P.
 XX PR 22-JAN-1998; 98US-072134P.
 XX PR 30-JAN-1998; 98US-073095P.
 XX PR 18-FEB-1998; 98US-075038P.
 XX PR 30-MAR-2000; 2000US-0539330.
 XX PR 23-NOV-1998; 98US-0197886.
 XX (JACO/) JACOBS K.
 PA (MCCO/) MCCOY J M.
 PA (LAVA/) LAVALLIE E R.
 PA (COLL/) COLLINS-RACIE L A.
 PA (EVAN/) EVANS C.
 PA (MERB/) MERBERG D.
 PA (TREA/) TREACY M.
 PA (AGOS/) AGOSTINO M J.
 PA (STEI/) STEININGER R J.

PA (SPAU/) SPAULDING V.
PA (WONG/) WONG G G.
PA (CLAR/) CLARK H.
PA (FECH/) FECHTEL K.
XX Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V;
PI Wong GG, Clark H, Fechtel K;
XX
DR WPI: 2002-040725/05.
DR N-PSDB; ABA90890.
XX
PT New secreted proteins and encoding polynucleotides, useful in gene
PT therapies, particularly for preventing or treating autoimmune
PT disorders, cancer, graft-versus-host disease, wound, osteoporosis,
PT stroke or inflammations -
XX
PS Disclosure; Page 197-198; 349pp; English.
XX
CC The invention relates to isolated polynucleotides (ABA90876-ABA90968 and
CC ABA90980) and encoded proteins (ABBS5598-ABBS5800), especially
CC polynucleotides SEQ ID NO 1 (ABA90876) and SEQ ID NO 19 (ABA90885) and
CC proteins SEQ ID NO 2 (ABBS5598) and SEQ ID NO 20 (ABBS5707) contained in
CC clones bd306-7 and y8-1 respectively and the clones bd306-7 and y8-1
CC are deposited with the American Type Culture Collection (ATCC) with
CC accession number 98599. The polynucleotides and encoded polypeptides have
CC cytotatic, anti-inflammatory, immunomodulator, vulnerary,
CC neuroprotective, activin, inhibin, chemotactic, haemostatic, thrombolytic
CC and anti-inflammatory activity and acting as cytokine modulators,
CC haematopoiesis regulators, tissue growth modulators and/or cadherin
CC suppressors. The polypeptides and polynucleotides are useful in gene
CC therapies, particularly for preventing, treating or ameliorating any of
CC the following diseases: immune deficiency and disorders; e.g. bacterial
CC or fungal infections, autoimmune disorders, cancer, systemic lupus
CC erythematosus or graft-versus-host disease; myeloid or lymphoid cell
CC deficiencies; wound, burns, incisions and ulcers, osteoporosis or
CC osteoarthritis; central and peripheral nervous system diseases and
CC neuropathies, e.g. Alzheimer's, Parkinson's disease, Huntington's
CC disease, amyotrophic lateral sclerosis or Shy-Drager syndrome;
CC haemophilia, cardiac infarction or stroke; inflammations, shock, sepsis
CC or systemic inflammatory response syndrome, ischaemia-reperfusion
CC injury, endotoxin lethality, arthritis, inflammatory bowel disease or
CC Crohn's disease; or tumours or cancers, pemphigus vulgaris or pemphigus
CC foliaceus.
XX
SQ Sequence 186 AA;
Query Match 2.0%; Score 8; DB 23; Length 186;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 81 SSGASTAT 88
Db 39 SSGASTAT 46
|||||||
RESULT 29
ABB71446
ID ABB71446 standard; Protein; 217 AA.
XX
AC ABB71446;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 41130.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX

PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI: 2001-656860/75.
DR N-PSDB; ABL15549.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 41130; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABBS7737-ABBS72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 217 AA;
Query Match 2.0%; Score 8; DB 22; Length 217;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 341 LQQLAAAV 348
|||||||
Db 33 LQQLAAAV 40
|||||||
RESULT 30
AAU58834
ID AAU58834 standard; Protein; 242 AA.
XX
AC AAU58834;
XX
DT 13-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #19730.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
DT 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX

XX
FI Hazelbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS56165.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 13899; Silpp; English.

The invention relates to antisense inhibitors of genes essential to prokaryotic cellular proliferation, their use in identifying the genes, their use in the discovery of novel antibiotics, the essential genes themselves and the encoded proteins. The prokaryotes used are Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence represents an essential prokaryotic cellular proliferation protein.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

SQ Sequence 254 AA; Query Match 2.0%; Score 8; DB 22; Length 254; Best Local Similarity 100.0%; Pred. No. 71; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 359 REAIQRLE 366
Dp 52 REAIQRLE 59

RESULT 32
AAR74043
ID AAR74043 standard: protein: 272 AA

AA AAR74043

DT 25-MAR-2003 (updated)

XX
 DT
 51-DEC-70
 (FIVE) ENCL

XX Rice carbonic-anhydrase.
XX
KW Rice: carbonic-anhydrase: carbon-fixation: monocotyledon: cereal:
DE
XX

XX
XX

XXXXXXXXXXXX

US
XX

PN WO9511979-A1.
YY

PD 04-MAY-1995.
XX

PP 27-OCT-1994; 94WO-JP01814.
yy

PR 29-OCT-1993; 93JP-0294278.

PA (NISB) JAPAN TOBACCO INC.

PI Burnell NJ. Suzuki S:

XX
DP WDT: 1005-170071/23

DR N-PSDB; AAQ92067.

XX Cloned DNA coding for monocotyledon carbonic anhydrase - is used
PT to improve carbon fixation in monocotyledonous plants.
XX
XX Claim 8; Page 17-18; 42pp; Japanese.
XX
CC The cDNA sequence given in AAQ2067 was obtd. by screening and
CC cloning a cDNA library generated from rice RNA, and encodes
CC carbonic-anhydrase.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 272 AA;

Query Match 2.0%; Score 8; DB 16; Length 272;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 AAPVAPAA 107
Db 63 AAPVAPAA 70
|||||

RESULT 33
AAG20687
ID AAG20687 standard; Protein; 309 AA.
XX
AC AAG20687;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 22973.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132048.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.

```
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.0%; Score 8; DB 21; Length 309;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 TLKGTNFE 14
Db 295 TLKGTNFE 302

RESULT 34
AAG38007
ID AAG38007 standard; Protein; 309 AA.
XX AAG38007;
AC AAG38007;
XX 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 46824.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 46824.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
```

PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149910.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

2.0%; Score 8; DB 21; Length 309;

Best Local Similarity 100.0%; Pred. No. 85;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 TLKGTNFE 14
|||||||
Db 295 TLKGTNFE 302

RESULT 35
AAG20686
ID AAG20686 standard; Protein; 316 AA.
XX AC AAG20686;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 22972.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.

```
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161559.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match      2.0%; Score 8; DB 21; Length 316;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 TLKGTNFE 14
Db 302 TLKGTNFE 309
```

RESULT 36

AAG38006

ID AAG38006 standard; Protein; 316 AA.

XX

AC AAG38006;

XX

DT 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 46823.

XX

KW Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.

PR

PR 05-MAR-1999; 99US-0123180.

PR

PR 09-MAR-1999; 99US-0123548.

PR

PR 23-MAR-1999; 99US-0125788.

PR

PR 25-MAR-1999; 99US-0126264.

PR

PR 29-MAR-1999; 99US-0126785.

PR

PR 01-APR-1999; 99US-0127462.

PR

PR 08-APR-1999; 99US-0128714.

PR

PR 16-APR-1999; 99US-0129845.

PR

PR 19-APR-1999; 99US-0130077.

PR

PR 21-APR-1999; 99US-0130449.

PR

PR 23-APR-1999; 99US-0130510.

PR

PR 28-APR-1999; 99US-0131449.

PR

PR 30-APR-1999; 99US-0132048.

PR

PR 30-APR-1999; 99US-0132407.

PR

PR 04-MAY-1999; 99US-0132484.

PR

PR 05-MAY-1999; 99US-0132485.

PR

PR 06-MAY-1999; 99US-0132486.

PR

PR 07-MAY-1999; 99US-0132487.

PR

PR 07-MAY-1999; 99US-0132863.

PR

PR 11-MAY-1999; 99US-0134256.

PR

PR 14-MAY-1999; 99US-0134218.

PR

PR 14-MAY-1999; 99US-0134219.

PR

PR 14-MAY-1999; 99US-0134221.

PR

PR 18-MAY-1999; 99US-0134768.

PR

PR 19-MAY-1999; 99US-0134941.

PR

PR 20-MAY-1999; 99US-0135124.

PR

PR 21-MAY-1999; 99US-0135353.

PR

PR 24-MAY-1999; 99US-0135629.

PR

PR 25-MAY-1999; 99US-0136021.

PR

PR 27-MAY-1999; 99US-0136392.

PR

PR 28-MAY-1999; 99US-0136782.

PR

PR 01-JUN-1999; 99US-0137222.

PR

PR 03-JUN-1999; 99US-0137528.

PR

PR 04-JUN-1999; 99US-0137502.

PR

PR 07-JUN-1999; 99US-0137724.

PR

PR 08-JUN-1999; 99US-0138094.

PR

PR 10-JUN-1999; 99US-0138540.

PR

PR 10-JUN-1999; 99US-0138847.

PR

PR 14-JUN-1999; 99US-0139119.

PR

PR 16-JUN-1999; 99US-0139452.

PR

PR 16-JUN-1999; 99US-0139453.

PR

PR 17-JUN-1999; 99US-0139492.

PR

PR 18-JUN-1999; 99US-0139454.

PR

PR 18-JUN-1999; 99US-0139455.

PR

PR 18-JUN-1999; 99US-0139456.

PR

PR 18-JUN-1999; 99US-0139457.

PR

PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.0%; Score 8; DB 21; Length 316;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 TLKGTNFE 14
 |||||
Db 302 TLKGTNFE 309

RESULT 37

AAG18084
ID AAG18084 standard; Protein; 338 AA.
XX
AC AAG18084;
XX

PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159322.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 2.0%; Score 8; DB 21; Length 339;
 Best Local Similarity 100.0%; Pred. No. 93;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 74 LSKAKASS 81
 |||||
 Db 11 LSKAKASS 18

RESULT 38

AAG20685
 ID AAG20685 standard; Protein; 342 AA.

XX AAG20685;

XX DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 22971.

XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; Genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX OS Arabidopsis thaliana.
 XX PN EP1033405-A2.
 XX PD 06-SEP-2000.
 XX PP 25-FEB-2000; 2000EP-0301439.
 XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.

```
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 27-JUL-1999; 99US-0145319.
PR 28-JUL-1999; 99US-0145351.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
```

```
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 2.0%; Score 8; DB 21; Length 342;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 TLKGTNFE 14
|||
Db 328 TLKGTNFE 335

RESULT 39

AAG38005
ID AAG38005 standard; Protein; 342 AA.
XX
AC AAG38005;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 46822.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX

PF 25-FEB-2000; 2000EP-0301439.
XX 99US-0121825.
PR 05-FEB-1999; 99US-0123180.
PR 05-MAR-1999; 99US-0123548.
PR 09-MAR-1999; 99US-0125788.
PR 23-MAR-1999; 99US-0126264.
PR 25-MAR-1999; 99US-0126785.
PR 29-MAR-1999; 99US-0127462.
PR 01-APR-1999; 99US-0128234.
PR 06-APR-1999; 99US-0128714.
PR 08-APR-1999; 99US-0129845.
PR 16-APR-1999; 99US-0130077.
PR 19-APR-1999; 99US-0130449.
PR 21-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 23-APR-1999; 99US-0131449.
PR 28-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 16-MAY-1999; 99US-0134768.
PR 18-MAY-1999; 99US-0134941.
PR 19-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0135353.
PR 21-MAY-1999; 99US-0135629.
PR 24-MAY-1999; 99US-0136021.
PR 25-MAY-1999; 99US-0136392.
PR 27-MAY-1999; 99US-0136782.
PR 28-MAY-1999; 99US-0137222.
PR 01-JUN-1999; 99US-0137528.
PR 03-JUN-1999; 99US-0137502.
PR 04-JUN-1999; 99US-0137724.
PR 07-JUN-1999; 99US-0138094.
PR 08-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 10-JUN-1999; 99US-0139119.
PR 14-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.

PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161320.
 PR 28-OCT-1999; 99US-0161392.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.
 PR 29-OCT-1999; 99US-0162143.
 PR 29-OCT-1999; 99US-0162144.

Query Match 2.0%; Score 8; DB 21; Length 342;
 Best Local Similarity 100.0%; Pred. No. 94;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 TLKGTNFE 14
 |||||
 DB 328 TLKGTNFE 335

RESULT 40
 ABP80988
 ID ABP80988 standard; Protein; 423 AA.

XX AC ABP80988;
 XX DT 07-MAR-2003 (first entry)
 XX DE N. gonorrhoeae amino acid sequence SEQ ID 8506.
 XX KW Antibacterial; infection; vaccine; gene therapy.
 XX OS Neisseria gonorrhoeae.

XX PN WO200279243-A2.
 XX PD 10-OCT-2002.
 XX PF 12-FEB-2002; 2002WO-IB02069.
 XX PR 12-FEB-2001; 2001GB-0003424.
 XX PA (CHIR-) CHIRON SPA.

XX PI Fontana MR, Pizza M, Masignani V, Monaci E;
 XX WPI; 2003-058415/05.

DR N-PSDB; ABZ41958.

XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
 PT medicament for treating or preventing N. gonorrhoeae infection -
 XX Disclosure; Page 807; 815pp; English.

XX The present invention relates to proteins from Neisseria gonorrhoeae.
 CC Also disclosed are the nucleic acid molecules encoding the proteins and
 CC antibodies that specifically bind to the proteins. The composition
 CC comprising the protein, nucleic acid or antibody is useful for the
 CC manufacture of a medicament for treating or preventing N. gonorrhoeae
 CC infection, this may be in the form of a vaccine or gene therapy.
 CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
 CC molecules of the invention.

SQ Sequence 423 AA;

Query Match 2.0%; Score 8; DB 24; Length 423;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 343 QLAAAVPQ 350
 |||||
 DB 28 QLAAAVPQ 35

RESULT 41

AMG48742
 ID AAG48742 standard; Protein; 536 AA.

XX AC AAG48742;
 XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 61585.
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.
 XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.
 XX PR 25-FEB-1999; 99US-0121825.
 XX PR 05-MAR-1999; 99US-0123180.
 XX PR 09-MAR-1999; 99US-0123548.
 XX PR 23-MAR-1999; 99US-0125788.
 XX PR 25-MAR-1999; 99US-0126264.
 XX PR 29-MAR-1999; 99US-0126785.
 XX PR 01-APR-1999; 99US-0127462.
 XX PR 06-APR-1999; 99US-0128234.
 XX PR 08-APR-1999; 99US-0128714.
 XX PR 16-APR-1999; 99US-0129845.
 XX PR 19-APR-1999; 99US-0130077.
 XX PR 21-APR-1999; 99US-0130449.
 XX PR 23-APR-1999; 99US-0130510.
 XX PR 23-APR-1999; 99US-0130891.
 XX PR 28-APR-1999; 99US-0131449.
 XX PR 30-APR-1999; 99US-0132048.
 XX PR 30-APR-1999; 99US-0132407.
 XX PR 04-MAY-1999; 99US-0132484.
 XX PR 05-MAY-1999; 99US-0132485.
 XX PR 06-MAY-1999; 99US-0132486.
 XX PR 06-MAY-1999; 99US-0132487.
 XX PR 07-MAY-1999; 99US-0132863.
 XX PR 11-MAY-1999; 99US-0134256.
 XX PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219;
PR 14-MAY-1999; 99US-0134221;
PR 14-MAY-1999; 99US-0134370;
PR 18-MAY-1999; 99US-0134768;
PR 19-MAY-1999; 99US-0134941;
PR 20-MAY-1999; 99US-0135124;
PR 21-MAY-1999; 99US-0135353;
PR 24-MAY-1999; 99US-0135629;
PR 25-MAY-1999; 99US-0136021;
PR 27-MAY-1999; 99US-0136392;
PR 28-MAY-1999; 99US-0136782;
PR 01-JUN-1999; 99US-0137222;
PR 03-JUN-1999; 99US-0137528;
PR 04-JUN-1999; 99US-0137502;
PR 07-JUN-1999; 99US-0137724;
PR 08-JUN-1999; 99US-0138094;
PR 10-JUN-1999; 99US-0138540;
PR 10-JUN-1999; 99US-0138847;
PR 14-JUN-1999; 99US-0139119;
PR 16-JUN-1999; 99US-0139452;
PR 16-JUN-1999; 99US-0139453;
PR 17-JUN-1999; 99US-0139492;
PR 18-JUN-1999; 99US-0139454;
PR 18-JUN-1999; 99US-0139455;
PR 18-JUN-1999; 99US-0139456;
PR 18-JUN-1999; 99US-0139457;
PR 18-JUN-1999; 99US-0139458;
PR 18-JUN-1999; 99US-0139459;
PR 18-JUN-1999; 99US-0139460;
PR 18-JUN-1999; 99US-0139461;
PR 18-JUN-1999; 99US-0139462;
PR 18-JUN-1999; 99US-0139463;
PR 18-JUN-1999; 99US-0139750;
PR 18-JUN-1999; 99US-0139763;
PR 21-JUN-1999; 99US-0139817;
PR 22-JUN-1999; 99US-0139899;
PR 23-JUN-1999; 99US-0140354;
PR 23-JUN-1999; 99US-0140355;
PR 24-JUN-1999; 99US-0140695;
PR 28-JUN-1999; 99US-0140823;
PR 29-JUN-1999; 99US-0140991;
PR 30-JUN-1999; 99US-0141287;
PR 01-JUL-1999; 99US-0141842;
PR 01-JUL-1999; 99US-0142154;
PR 02-JUL-1999; 99US-0142055;
PR 06-JUL-1999; 99US-0142390;
PR 08-JUL-1999; 99US-0142803;
PR 09-JUL-1999; 99US-0142920;
PR 12-JUL-1999; 99US-0142977;
PR 13-JUL-1999; 99US-0143542;
PR 14-JUL-1999; 99US-0143624;
PR 15-JUL-1999; 99US-0144005;
PR 16-JUL-1999; 99US-0144005;
PR 16-JUL-1999; 99US-0144085;
PR 19-JUL-1999; 99US-0144325;
PR 19-JUL-1999; 99US-0144331;
PR 19-JUL-1999; 99US-0144332;
PR 19-JUL-1999; 99US-0144333;
PR 19-JUL-1999; 99US-0144334;
PR 19-JUL-1999; 99US-0144335;
PR 20-JUL-1999; 99US-0144352;
PR 20-JUL-1999; 99US-0144632;
PR 20-JUL-1999; 99US-0144884;
PR 21-JUL-1999; 99US-0144814;
PR 21-JUL-1999; 99US-0145086;
PR 21-JUL-1999; 99US-0145088;
PR 22-JUL-1999; 99US-0145085;
PR 22-JUL-1999; 99US-0145087;
PR 22-JUL-1999; 99US-0145089;
PR 23-JUL-1999; 99US-0145192;
PR 23-JUL-1999; 99US-0145145;
PR 23-JUL-1999; 99US-0145210;
PR 23-JUL-1999; 99US-0145224;
PR 26-JUL-1999; 99US-0145276;
PR 27-JUL-1999; 99US-0145913;
PR 27-JUL-1999; 99US-0145918;
PR 27-JUL-1999; 99US-0145919;
PR 28-JUL-1999; 99US-0145951;
PR 02-AUG-1999; 99US-0146386;
PR 02-AUG-1999; 99US-0146388;
PR 02-AUG-1999; 99US-0146389;
PR 03-AUG-1999; 99US-0147038;
PR 04-AUG-1999; 99US-0147204;
PR 04-AUG-1999; 99US-0147302;
PR 05-AUG-1999; 99US-0147192;
PR 05-AUG-1999; 99US-0147260;
PR 06-AUG-1999; 99US-0147303;
PR 06-AUG-1999; 99US-0147416;
PR 09-AUG-1999; 99US-0147493;
PR 09-AUG-1999; 99US-0147935;
PR 10-AUG-1999; 99US-0148171;
PR 11-AUG-1999; 99US-0148319;
PR 12-AUG-1999; 99US-0148341;
PR 13-AUG-1999; 99US-0148565;
PR 13-AUG-1999; 99US-0148684;
PR 16-AUG-1999; 99US-0149368;
PR 17-AUG-1999; 99US-0149175;
PR 18-AUG-1999; 99US-0149426;
PR 20-AUG-1999; 99US-0149722;
PR 20-AUG-1999; 99US-0149723;
PR 20-AUG-1999; 99US-0149929;
PR 23-AUG-1999; 99US-0149902;
PR 23-AUG-1999; 99US-0149930;
PR 25-AUG-1999; 99US-0150566;
PR 26-AUG-1999; 99US-0150884;
PR 27-AUG-1999; 99US-0151065;
PR 27-AUG-1999; 99US-0151066;
PR 27-AUG-1999; 99US-0151080;
PR 30-AUG-1999; 99US-0151303;
PR 31-AUG-1999; 99US-0151438;
PR 01-SEP-1999; 99US-0151930;
PR 07-SEP-1999; 99US-0152363;
PR 10-SEP-1999; 99US-0153070;
PR 13-SEP-1999; 99US-0153758;
PR 15-SEP-1999; 99US-0154018;
PR 16-SEP-1999; 99US-0154039;
PR 20-SEP-1999; 99US-0154779;
PR 22-SEP-1999; 99US-0155139;
PR 23-SEP-1999; 99US-0155486;
PR 24-SEP-1999; 99US-0155659;
PR 28-SEP-1999; 99US-0156458;
PR 29-SEP-1999; 99US-0156596;
PR 04-OCT-1999; 99US-0157117;
PR 05-OCT-1999; 99US-0157753;
PR 06-OCT-1999; 99US-0157865;
PR 07-OCT-1999; 99US-0158029;
PR 08-OCT-1999; 99US-0158232;
PR 12-OCT-1999; 99US-0158369;
PR 13-OCT-1999; 99US-0159293;
PR 13-OCT-1999; 99US-0159294;
PR 13-OCT-1999; 99US-0159295;
PR 14-OCT-1999; 99US-0159329;
PR 14-OCT-1999; 99US-0159330;
PR 14-OCT-1999; 99US-0159331;
PR 14-OCT-1999; 99US-0159637;
PR 18-OCT-1999; 99US-0159638;
PR 18-OCT-1999; 99US-0159584;
PR 21-OCT-1999; 99US-0160741;
PR 21-OCT-1999; 99US-0160767;
PR 21-OCT-1999; 99US-0160768;
PR 21-OCT-1999; 99US-0160770;
PR 21-OCT-1999; 99US-0160814;
PR 21-OCT-1999; 99US-0160815;
PR 22-OCT-1999; 99US-0160980;
PR 22-OCT-1999; 99US-0160981;
PR 22-OCT-1999; 99US-0160989;

PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 2.0%; Score 8; DB 21; Length 536;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 LSKAKASS 81
 DB 10 LSKAKASS 17
 |||||

RESULT 42
 AAW89794
 ID AAW89794 standard; Protein; 540 AA.
 XX
 AC AAW89794;
 XX
 DT 16-MAR-1999 (first entry)
 XX
 DE Staphylococcus aureus protein SEQ ID #5242.

XX Computer readable medium; vaccine; S.aureus infection; immunodetection;
 KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KW skin infection; surgical wound infection; scalded skin syndrome;
 KW toxic shock syndrome.
 XX
 OS Staphylococcus aureus.
 XX
 PN EP786519-A2.
 XX
 PD 30-JUL-1997.

XX 07-JAN-1997; 97EP-0100117.
 XX
 PR 05-JAN-1996; 96US-0009861.
 XX
 PR (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
 PI Rosen CA;
 PI
 XX WPI; 1997-374922/35.
 XX
 XX Polynucleotide(s) and proteins derived from Staphylococcus aureus -
 PT stored on computer readable medium and used in the production of
 PT anti-S.aureus vaccines
 XX
 XX Claim 23; Page 3244-3246; 3271pp; English.

XX This sequence represents a Staphylococcus aureus protein sequence of the
 CC invention. The DNA sequences encoding the S.aureus proteins are recorded
 CC on a computer readable medium, preferably selected from a floppy or hard
 CC disk, random access memory (RAM), read-only memory (ROM) or CD-ROM.
 CC Homology searches using the S.aureus DNA sequences allows putative
 CC functions to be assigned so that protein-encoding or regulatory regions
 CC of commercial, therapeutic or industrial importance can be obtained.
 CC Specifically, sequences which are likely to encode antigens have been
 CC identified and these polypeptides can be used in a vaccine composition
 CC against S.aureus infection. The polypeptides can also be used in a kit
 CC for the immunodetection of S.aureus in a sample. S.aureus is implicated
 CC in numerous human diseases, including cellulitis, eyelid infections, food
 CC poisoning, osteomyelitis, skin and surgical wound infections, scalded
 CC skin syndrome, toxic shock syndrome, etc. Organisms transformed with the
 CC DNA sequences can be used for recombinant production of the polypeptides.

CC The new DNA sequences (and their fragments) are useful as primers or
 CC probes for isolating homologues of any of the 5191 S.aureus DNA sequences
 CC contained on the computer readable medium.

XX Sequence 540 AA;

Query Match 2.0%; Score 8; DB 18; Length 540;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 AQPAPVA 104
 DB 175 AQPAPVA 182
 |||||

RESULT 43
 AAW89777
 ID AAW89777 standard; Protein; 540 AA.
 XX
 AC AAW89777;
 XX
 DT 16-MAR-1999 (first entry)
 XX
 DE Staphylococcus aureus protein SEQ ID #5225.

XX Computer readable medium; vaccine; S.aureus infection; immunodetection;
 KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KW skin infection; surgical wound infection; scalded skin syndrome;
 KW toxic shock syndrome.
 XX
 OS Staphylococcus aureus.
 XX
 PN EP786519-A2.
 XX
 PD 30-JUL-1997.
 XX
 XX 07-JAN-1997; 97EP-0100117.
 XX
 PR 05-JAN-1996; 96US-0009861.
 XX
 PR (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
 PI Rosen CA;
 PI
 XX WPI; 1997-374922/35.

XX Polynucleotide(s) and proteins derived from Staphylococcus aureus -
 PT stored on computer readable medium and used in the production of
 PT anti-S.aureus vaccines
 XX
 XX Claim 23; Page 3221-3222; 3271pp; English.

XX This sequence represents a Staphylococcus aureus protein sequence of the
 CC invention. The DNA sequences encoding the S.aureus proteins are recorded
 CC on a computer readable medium, preferably selected from a floppy or hard
 CC disk, random access memory (RAM), read-only memory (ROM) or CD-ROM.
 CC Homology searches using the S.aureus DNA sequences allows putative
 CC functions to be assigned so that protein-encoding or regulatory regions

CC of commercial, therapeutic or industrial importance can be obtained.
 CC Specifically, sequences which are likely to encode antigens have been
 CC identified and these polypeptides can be used in a vaccine composition
 CC against *S.aureus* infection. The polypeptides can also be used in a kit
 CC for the immunodetection of *S.aureus* in a sample. *S.aureus* is implicated
 CC in numerous human diseases, including cellulitis, eyelid infections, food
 CC poisoning, osteomyelitis, skin and surgical wound infections, scalded
 CC skin syndrome, toxic shock syndrome, etc. Organisms transformed with the
 CC DNA sequences can be used for recombinant production of the polypeptides.
 CC The new DNA sequences (and their fragments) are useful as primers or
 CC probes for isolating homologues of any of the 5191 *S.aureus* DNA sequences
 CC contained on the computer readable medium.

XX Sequence 540 AA;

Query Match 2.0%; Score 8; DB 18; Length 540;

Best Local Similarity 100.0%; Pred.No.1.5e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 AQPAPVA 104

Db 175 AQPAPVA 182

RESULT 44

AA99408

ID AA99408 standard; Protein; 596 AA.

AC AA99408;

XX

XX

DT 08-AUG-2000 (first entry)

XX

XX

DE Human PRO1342 (UNQ697) amino acid sequence SEQ ID NO:243.

XX

XX

KW Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;

KW transmembrane; secretion; immunoadhesion; pharmaceutical; screening.

XX

XX

PN WO200012708-A2.

XX

XX

PD 09-MAR-2000.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PR 17-SEP-1998; 98US-0100684.
 PR 17-SEP-1998; 98US-0100710.
 PR 17-SEP-1998; 98US-0100711.
 PR 17-SEP-1998; 98US-0100919.
 PR 17-SEP-1998; 98US-0100930.
 PR 18-SEP-1998; 98US-0100848.
 PR 18-SEP-1998; 98US-0100849.
 PR 18-SEP-1998; 98US-0101014.
 PR 18-SEP-1998; 98US-0101068.
 PR 18-SEP-1998; 98US-0101071.
 PR 18-SEP-1998; 98US-0101279.
 PR 23-SEP-1998; 98US-0101471.
 PR 23-SEP-1998; 98US-0101472.
 PR 23-SEP-1998; 98US-0101474.
 PR 23-SEP-1998; 98US-0101475.
 PR 23-SEP-1998; 98US-0101476.
 PR 23-SEP-1998; 98US-0101477.
 PR 23-SEP-1998; 98US-0101479.
 PR 24-SEP-1998; 98US-0101738.
 PR 24-SEP-1998; 98US-0101741.
 PR 24-SEP-1998; 98US-0101743.
 PR 24-SEP-1998; 98US-0101915.
 PR 24-SEP-1998; 98US-0101916.
 PR 29-SEP-1998; 98US-0102207.
 PR 29-SEP-1998; 98US-0102240.
 PR 29-SEP-1998; 98US-0102307.
 PR 29-SEP-1998; 98US-0102330.
 PR 30-SEP-1998; 98US-0102331.
 PR 30-SEP-1998; 98US-0102484.
 PR 30-SEP-1998; 98US-0102487.
 PR 30-SEP-1998; 98US-0102570.
 PR 01-OCT-1998; 98US-0102571.
 PR 01-OCT-1998; 98US-0102684.
 PR 01-OCT-1998; 98US-0102687.
 PR 02-OCT-1998; 98US-0102965.
 PR 06-OCT-1998; 98US-0103258.
 PR 06-OCT-1998; 98US-0103449.
 PR 07-OCT-1998; 98US-0103314.
 PR 07-OCT-1998; 98US-0103315.
 PR 07-OCT-1998; 98US-0103328.
 PR 07-OCT-1998; 98US-0103395.
 PR 07-OCT-1998; 98US-0103396.
 PR 07-OCT-1998; 98US-0103401.
 PR 08-OCT-1998; 98US-0103633.
 PR 08-OCT-1998; 98US-0103678.
 PR 08-OCT-1998; 98US-0103679.
 PR 08-OCT-1998; 98US-0103711.
 PR 14-OCT-1998; 98US-0104257.
 PR 20-OCT-1998; 98US-0104987.
 PR 20-OCT-1998; 98US-0105000.
 PR 20-OCT-1998; 98US-0105002.
 PR 21-OCT-1998; 98US-0105104.
 PR 22-OCT-1998; 98US-0105169.
 PR 22-OCT-1998; 98US-0105266.
 PR 26-OCT-1998; 98US-0105693.
 PR 26-OCT-1998; 98US-0105694.
 PR 27-OCT-1998; 98US-0105807.
 PR 27-OCT-1998; 98US-0105881.
 PR 27-OCT-1998; 98US-0105882.
 PR 28-OCT-1998; 98US-0106023.
 PR 28-OCT-1998; 98US-0106029.
 PR 28-OCT-1998; 98US-0106030.
 PR 28-OCT-1998; 98US-0106032.
 PR 28-OCT-1998; 98US-0106033.
 PR 29-OCT-1998; 98US-0106178.
 PR 29-OCT-1998; 98US-0106248.
 PR 29-OCT-1998; 98US-0106384.
 PR 29-OCT-1998; 98US-0106500.
 PR 30-OCT-1998; 98US-0106464.
 PR 03-NOV-1998; 98US-0106855.
 PR 03-NOV-1998; 98US-0106902.
 PR 03-NOV-1998; 98US-0106905.

PR 03-NOV-1998; 98US-0106919.
PR 03-NOV-1998; 98US-0106932.
PR 03-NOV-1998; 98US-0106934.
PR 10-NOV-1998; 98US-0107783.
PR 17-NOV-1998; 98US-0108775.
PR 17-NOV-1998; 98US-0108779.
PR 17-NOV-1998; 98US-0108787.
PR 17-NOV-1998; 98US-0108788.
PR 17-NOV-1998; 98US-0108801.
PR 17-NOV-1998; 98US-0108802.
PR 17-NOV-1998; 98US-0108806.
PR 17-NOV-1998; 98US-0108807.
PR 17-NOV-1998; 98US-0108867.
PR 17-NOV-1998; 98US-0108925.
PR 18-NOV-1998; 98US-0108848.
PR 18-NOV-1998; 98US-0108849.
PR 18-NOV-1998; 98US-0108850.
PR 18-NOV-1998; 98US-0108851.
PR 18-NOV-1998; 98US-0108852.
PR 18-NOV-1998; 98US-0108858.
PR 18-NOV-1998; 98US-0108904.
XX (GETH) GENENTECH INC.
XX Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;
XX WPI; 2000-237871/20.
XX N-PSDB; AAA37090.
XX New mammalian DNA sequences encoding transmembrane, receptor or
PT secreted PRO polypeptides, useful for screening of potential peptide or
PT small molecule inhibitors of the relevant receptor/ligand interactions
XX Claim 12; Fig 138; 773pp; English.
XX AAA37022 to AAA37144 encode the new isolated human transmembrane,
CC receptor or secreted PRO polypeptides given in AAY99340 to AAY99462. The
CC transmembrane and receptor PRO proteins can be used for screening of
CC potential peptide or small molecule inhibitors of the relevant
CC receptor/ligand interactions. The polypeptides and nucleotide sequences
CC encoding then have various industrial applications, including uses as
CC pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent
CC PCR primers and hybridisation probes used in the isolation of the PRO
CC polypeptides from the present invention.
XX Sequence 596 AA;
SQ
Query Match 2.0%; Score 8; DB 21; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 81 SSGASTAT 88
Db 39 SSGASTAT 46
|||||
RESULT 45
AAU29178
ID AAU29178 standard; Protein; 596 AA.
XX AAU29178;
XX AC
XX AAU29178;
DT 18-DEC-2001 (first entry)
XX Human PRO polypeptide sequence #155.
XX PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
XX Homo sapiens.
XX OS
XX

PN WO200168849-A2.
XX 20-SEP-2001.
XX 28-FEB-2001; 2001WO-US06520.
XX 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 03-MAR-2000; 2000US-187202P.
PR 06-MAR-2000; 2000US-186968P.
PR 14-MAR-2000; 2000US-189320P.
PR 15-MAR-2000; 2000US-189328P.
PR 15-MAR-2000; 2000WO-US06884.
PR 21-MAR-2000; 2000US-190828P.
PR 21-MAR-2000; 2000US-191007P.
PR 21-MAR-2000; 2000US-191048P.
PR 21-MAR-2000; 2000US-191314P.
PR 28-MAR-2000; 2000US-192655P.
PR 29-MAR-2000; 2000US-193032P.
PR 29-MAR-2000; 2000US-193053P.
PR 30-MAR-2000; 2000WO-US08439.
PR 04-APR-2000; 2000US-194449P.
PR 04-APR-2000; 2000US-194647P.
PR 11-APR-2000; 2000US-195975P.
PR 11-APR-2000; 2000US-196000P.
PR 11-APR-2000; 2000US-196187P.
PR 11-APR-2000; 2000US-196690P.
PR 11-APR-2000; 2000US-196820P.
PR 18-APR-2000; 2000US-198121P.
PR 18-APR-2000; 2000US-198585P.
PR 25-APR-2000; 2000US-199397P.
PR 25-APR-2000; 2000US-199550P.
PR 25-APR-2000; 2000US-199654P.
PR 03-MAY-2000; 2000US-201516P.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 05-JUN-2000; 2000US-209832P.
PR 28-JUL-2000; 2000WO-US20710.
PR 22-AUG-2000; 2000US-0644848.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
XX (GETH) GENENTECH INC.
XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX WPI; 2001-602746/68.
XX N-PSDB; AAS46079.
XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the
PT presence of tumours, such as prostate and breast tumours, in mammals and
PT to screen for modulators of the compounds -
XX Claim 11; Fig 310; 774pp; English.
XX Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
CC The PRO polypeptides and their associated nucleic acids can be used to
CC detect the presence of a tumour in a mammal by comparing the level of
CC expression of a PRO polypeptide in a test sample of cells from the animal
CC and a control sample of normal cells, whereby a higher level of
CC expression in the test sample indicates the presence of a tumour in the
CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
CC and rabbits but are preferably human. The polypeptides can be used to
CC stimulate tumour necrosis factor (TNF) alpha release from human blood,
CC when contacted with it. A specific polypeptide can be used to stimulate
CC the proliferation or differentiation of chondrocyte cells. The PRO
CC proteins can be used to determine the presence of tumours and also
CC susceptibility to tumour development, particularly adrenal, lung, colon,

CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
 CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
 CC can be used for genetic analysis of individuals with genetic disorders.

XX Sequence 596 AA;

Query Match 2.0%; Score 8; DB 22; Length 596;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
 DB 39 SSGASTAT 46
 |||||

RESULT 46

AAB87575
 ID AAB87575 standard; Protein; 596 AA.

XX AAB87575;

DT 15-MAY-2001 (first entry)

DE Human PRO1342.

XX Human; PRO protein; mapping.

XX Homo sapiens.

PN WO200116318-A2.

XX 08-MAR-2001.

XX 24-AUG-2000; 2000WO-US23328.

PR 01-SEP-1999; 99WO-US20111.

PR 15-SEP-1999; 99WO-US21090.

PR 07-DEC-1999; 99US-0169495.

PR 09-DEC-1999; 99US-0170262.

PR 11-JAN-2000; 2000US-0175481.

PR 18-FEB-2000; 2000WO-US04341.

PR 22-FEB-2000; 2000WO-US04414.

PR 01-MAR-2000; 2000WO-US05601.

PR 03-MAR-2000; 2000US-0187202.

PR 25-APR-2000; 2000US-0199397.

PR 22-MAY-2000; 2000WO-US14042.

PR 05-JUN-2000; 2000US-0209832.

XX (GETH) GENENTECH INC.

PI Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;

PI Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;

XX WPI; 2001-183260/18.

DR N-PSDB; AAF92107.

XX Eighty four nucleic acids encoding PRO polypeptides, useful in

PT molecular biology, including use as hybridization probes, and in

PT chromosome and gene mapping. -

XX Claim 12; Fig 100; 278pp; English.

PS The present sequence is a human PRO polypeptide (secreted and

CC transmembrane). The PRO protein, and PRO agonists, PRO antagonists or

CC anti-PRO antibodies are useful for preparation of a medicament useful in

CC the treatment of a condition which is responsive to the PRO protein,

CC agonists, antagonists or anti-PRO antibodies. The PRO protein may also be

CC employed as molecular weight markers for protein electrophoresis. The PRO

CC coding sequence has applications in molecular biology, including use as

CC hybridisation probes, and in chromosome and gene mapping.

XX Sequence 596 AA;

Query Match

Best Local Similarity 100.0%; Score 8; DB 22; Length 596;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
 DB 39 SSGASTAT 46
 |||||

RESULT 47

AAB66157
 ID AAB66157 standard; protein; 596 AA.

XX AAB66157;

DT 02-APR-2001 (first entry)

DE Protein of the invention #69.

XX Secreted; transmembrane; gene therapy.

OS Unidentified.

PN WO200078961-A1.

PD 28-DEC-2000.

PF 18-FEB-2000; 2000WO-US04342.

PR 23-JUN-1999; 99US-0141037.

PR 26-JUL-1999; 99US-0144758.

PR 01-SEP-1999; 99WO-US20111.

PR 29-OCT-1999; 99US-0182506.

PR 30-NOV-1999; 99WO-US28313.

PR 16-DEC-1999; 99WO-US30095.

PR 05-JAN-2000; 2000WO-US00219.

XX 06-JAN-2000; 2000WO-US00376.

PA (GETH) GENENTECH INC.

PI Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;

PI Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;

PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D;

PI Watanabe CK, Williams PM, Wood WI;

XX WPI; 2001-071395/08.

DR Secreted and transmembrane proteins and nucleic acids designated PRO,

PT useful as hybridization probes, in chromosome and gene mapping and gene

PT therapy -

XX Claim 1; Fig 138; 787pp; English.

PS The present invention relates to secreted and transmembrane proteins.

CC These proteins and the DNA encoding them may be used as hybridization

CC probes, in chromosome and gene mapping and in the generation of

CC anti-sense RNA and DNA. They may also be used to generate either

CC transgenic animals or knockout animals which are in turn useful for

CC development and screening of therapeutically useful reagents.

CC The nucleic acids may also be used in gene therapy.

XX Sequence 596 AA;

Query Match

Best Local Similarity 100.0%; Score 8; DB 22; Length 596;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
 DB 39 SSGASTAT 46
 |||||

RESULT 48

ABG95900
 ID ARG95900 standard; Protein; 596 AA.
 AC ABG95900;
 DT 10-DEC-2002 (first entry)
 XX Human secreted/transmembrane protein PRO1342.
 DE Human; secreted protein; transmembrane protein; antirheumatic;
 XX antiarthritic; osteopathic; sports-related joint problem;
 KW articular cartilage defect; osteoarthritis; rheumatoid arthritis.
 KW Homo sapiens.
 OS
 XX US2002119130-A1.
 PN
 XX 29-AUG-2002.
 PD
 XX 06-DEC-2001; 2001US-0006867.
 PF
 XX 29-OCT-1997; 97US-063435P.
 PR 29-OCT-1997; 97US-064215P.
 PR 22-APR-1998; 98US-082797P.
 PR 29-APR-1998; 98US-083495P.
 PR 15-MAY-1998; 98US-085579P.
 PR 10-JUN-1998; 98US-088811P.
 PR 10-JUN-1998; 98US-088824P.
 PR 11-JUN-1998; 98US-088863P.
 PR 12-JUN-1998; 98US-089105P.
 PR 16-JUN-1998; 98US-089514P.
 PR 16-SEP-1998; 98WO-US19330.
 PR 08-MAR-1999; 99WO-US05028.
 PR 14-MAY-1999; 99WO-US10733.
 PR 02-JUN-1999; 99WO-US12252.
 PR 01-SEP-1999; 99WO-US20111.
 PR 15-SEP-1999; 99WO-US21090.
 PR 15-SEP-1999; 99WO-US21194.
 PR 22-DEC-1999; 99WO-US30720.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 18-FEB-2000; 2000WO-US04342.
 PR 30-MAR-2000; 2000WO-US08439.
 PR 22-MAY-2000; 2000WO-US14042.
 PR 02-JUN-2000; 2000WO-US15264.
 PR 23-AUG-2000; 2000WO-US23522.
 PR 24-AUG-2000; 2000WO-US23328.
 PR 10-NOV-2000; 2000WO-US30873.
 PR 01-DEC-2000; 2000WO-US32378.
 PR 20-DEC-2000; 2000WO-US34956.
 PR 28-FEB-2001; 2001WO-US06520.
 PR 20-JUN-2001; 2001WO-US19692.
 PR 29-JUN-2001; 2001WO-US21066.
 PR 09-JUL-2001; 2001WO-US21735.
 XX (GETH) GENENTECH INC.
 PA
 XX Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;
 PI Grimaldi JC, Gurney AL, Watanabe CK, Wood WI;
 PI WPI; 2002-731348/79.
 DR N-PSDB; ABS74427.
 DR
 XX New isolated secreted and transmembrane PRO polypeptide useful for
 PT modulating biological activity of a cell, or for treating
 PT sports-related joint problems, osteoarthritis or rheumatoid arthritis
 XX
 PS Claim 20; Fig 100; 399pp; English.
 XX
 CC The invention relates to an isolated secreted and transmembrane PRO

CC polypeptide having 80 % sequence identity to a sequence appearing
 CC as ABG95851-ABG95934 or their associated signal peptide, or a sequence of
 CC an extracellular domain of the proteins with their associated signal
 CC peptide or lacking its associated signal peptide. Also included are
 CC the nucleic acids encoding the proteins, vectors, host cells,
 CC fusion proteins and antibodies which specifically bind to the proteins.
 CC The proteins are useful for detecting a polypeptide designated as A, B, C
 CC or D in a sample suspected of containing an A, B, C or D polypeptide,
 CC by contacting the sample with a polypeptide designated as E, F, G, H or
 CC I (or vice versa) and determining the formation of a A/E, B/F, B/G, C/H
 CC or D/I polypeptide conjugate in the sample, where the formation of the
 CC conjugate is indicative of the presence of an A, B, C or D polypeptide
 CC in the sample, where A is a PRO10272 polypeptide, B is a PRO20110
 CC polypeptide, C is a PRO10096 polypeptide, D is a PRO19760 polypeptide,
 CC E is a PRO5801 polypeptide, F is a PRO1 polypeptide, G is a PRO20040
 CC polypeptide, H is a PRO20233 polypeptide and I is a PRO1990
 CC polypeptide. The sample comprises a cell suspected of expressing the A,
 CC B, C or D polypeptide. The E, F, G, H or I polypeptide is labeled with
 CC a detectable label or is attached to a solid support. The proteins are
 CC useful for linking a bioactive molecule to a cell expressing a
 CC polypeptide designated as A, B, C or D or E, F, G, H or I. The bioactive
 CC molecule is a toxin, a radiolabel or an antibody. The bioactive molecule
 CC causes death of the cell. A, B, C, D, E, F, G, H, or I, or antibodies
 CC against them are useful for modulating a biological activity of a cell
 CC expressing a polypeptide designated as A, B, C or D or E, F, G, H, or
 CC I. The cell is killed. The proteins are useful for identifying
 CC agonists or antagonists, for the preparation of a medicament useful in
 CC the treatment of a condition which is responsive to the proteins, as
 CC molecular weight markers for protein electrophoresis purposes, and as
 CC therapeutic agents for treating sports-related joint problems,
 CC articular cartilage defects, osteoarthritis or rheumatoid arthritis.
 CC Nucleic acids encoding the proteins are useful as hybridisation probes,
 CC in chromosome and gene mapping, in the generation of anti-sense RNA and
 CC DNA, for the preparation of the proteins, to generate transgenic or
 CC knockout animals which are useful in the development and screening of
 CC therapeutic useful reagents, for chromosome identification, and in gene
 CC therapy. The antibody is useful as a therapeutic agent, and in gene
 CC assay and for affinity purification of the protein from recombinant
 CC cell culture natural sources. The present sequence represents a novel
 CC secreted or transmembrane protein of the invention.

XX Sequence 596 AA;

Query Match 2.0%; Score 8; DB 23; Length 596;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
 |||||
 DB 39 SSGASTAT 46

RESULT 49

ABU71266
 ID ABU71266 standard; Protein; 596 AA.

XX AC ABU71266;

DT 10-JUN-2003 (first entry)

XX DE Human PRO1342 protein.

XX Human; PRO; secreted; transmembrane; cytotstatic; TNF-alpha; blood;
 KW tumour necrosis factor alpha release; chondrocyte cell; proliferation;
 KW differentiation; tumour; gene therapy.

XX OS Homo sapiens.

XX US2003036143-A1.

XX PD 20-FEB-2003.

XX PF 02-JUL-2002; 2002US-0187600.

XX PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US05520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 28-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085573P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088752P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090423P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.

PR 01-JUL-1998; 98US-0911359P.
 PR 01-JUL-1998; 98US-0911544P.
 PR 02-JUL-1998; 98US-091478P.
 PR 02-JUL-1998; 98US-091486P.
 PR 02-JUL-1998; 98US-091626P.
 PR 02-JUL-1998; 98US-091628P.
 PR 02-JUL-1998; 98US-091632P.
 PR 24-JUL-1998; 98US-094006P.
 PR 04-AUG-1998; 98US-095282P.
 PR 10-AUG-1998; 98US-095998P.
 PR 10-AUG-1998; 98US-096012P.
 PR 17-AUG-1998; 98US-096757P.
 PR 17-AUG-1998; 98US-096766P.
 PR 17-AUG-1998; 98US-096867P.
 PR 17-AUG-1998; 98US-096891P.
 PR 17-AUG-1998; 98US-096897P.
 PR 18-AUG-1998; 98US-096949P.
 PR 18-AUG-1998; 98US-096959P.
 PR 26-AUG-1998; 98US-097022P.
 PR 26-AUG-1998; 98US-097352P.
 PR 26-AUG-1998; 98US-097354P.
 PR 26-AUG-1998; 98US-097355P.
 PR 26-AUG-1998; 98US-097971P.
 PR 26-AUG-1998; 98US-097974P.
 PR 01-SEP-1998; 98US-098014P.
 PR 01-SEP-1998; 98US-098116P.
 PR 02-SEP-1998; 98US-098723P.
 PR 02-SEP-1998; 98US-098803P.
 PR 02-SEP-1998; 98US-098821P.
 PR 09-SEP-1998; 98US-098843P.
 PR 09-SEP-1998; 98US-099602P.
 PR 10-SEP-1998; 98US-099741P.
 PR 10-SEP-1998; 98US-099754P.
 PR 10-SEP-1998; 98US-099763P.
 PR 10-SEP-1998; 98US-099812P.

Query Match 2.0%; Score 8; DB 24; Length 596;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
 |||||
 Db 39 SSGASTAT 46

RESULT 50
 ABU71555
 ID ABU71555 standard; Protein; 596 AA.
 AC ABU71555;
 XX
 XX 10-JUN-2003 (first entry)
 XX Human secreted polypeptide PRO1342.
 DE
 XX Human; gene therapy; tumour; cancer.
 XX Homo sapiens.
 OS
 XX US2003013855-A1.
 PN
 XX 16-JAN-2003.
 PD
 XX 03-MAY-2002; 2002US-0063616.

XX 30-DEC-1998; 98KR-0062142.
 PR 08-MAR-1999; 99WO-US05028.
 PR 14-MAY-1999; 99WO-US10733.
 PR 30-DEC-1999; 99WO-US31274.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 01-MAR-2000; 2000WO-US05601.
 PR 02-MAR-2000; 2000WO-US05841.
 PR 21-MAR-2000; 2000WO-US07532.

PR 22-MAY-2000; 2000WO-US14042.
 PR 02-JUN-2000; 2000WO-US15264.
 PR 24-AUG-2000; 2000WO-US23328.
 PR 10-NOV-2000; 2000WO-US30873.
 PR 01-DEC-2000; 2000WO-US32678.
 PR 20-DEC-2000; 2000WO-US34956.
 PR 28-FEB-2001; 2001WO-US06520.
 PR 01-JUN-2001; 2001WO-US17800.
 PR 14-MAY-1999; 99US-0311832.
 PR 25-AUG-1999; 99US-0380137.
 PR 25-AUG-1999; 99US-0380138.
 PR 25-AUG-1999; 99US-0380139.
 PR 15-SEP-1999; 99US-0380142.
 PR 15-SEP-1999; 99US-0397342.
 PR 18-OCT-1999; 99US-0403297.
 PR 12-NOV-1999; 99US-0423844.
 PR 22-AUG-2000; 2000US-0644848.
 PR 18-SEP-2000; 2000US-0664610.
 PR 18-SEP-2000; 2000US-0665350.
 PR 08-NOV-2000; 2000US-0709238.
 PR 20-DEC-2000; 2000US-0747259.
 PR 22-MAR-2001; 2001US-0816744.
 PR 10-MAY-2001; 2001US-0854208.
 PR 30-MAY-2001; 2001US-0854280.
 PR 05-JUN-2001; 2001US-0870574.
 PR 05-JUN-2001; 2001US-0874503.
 PR 29-JUN-2001; 2001US-0869599.
 PR 18-JUL-2001; 2001US-0908827.
 PR 06-DEC-2001; 2001US-0006867.

(GETH) GENENTECH INC.

PA Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;
 PI Gramaldi JC, Gurney AL, Watanabe CK, Wood WI;
 XX

DR WPI; 2003-330485/31.
 DR N-PSDB; ACA58859.

XX New isolated antibody specifically binding a PRO polypeptide, useful
 PT for the preparation of a medicament for treating disorders with the
 PT aberrant expression or activity of the PRO polypeptide, such as tumor
 PT conditions and cancer -
 XX

PS Disclosure; Page 169-171; 406pp; English.

XX The invention relates to an antibody that binds to a polypeptide with a
 CC fully defined sequence given in the specification. The methods and
 CC compositions (containing antibodies that specifically bind a PRO
 CC polypeptide) of the present invention are useful for the preparation of a
 CC medicament for the treatment of disorders associated with the aberrant
 CC expression or activity of the PRO polypeptide, such as tumour conditions
 CC and cancer. They can also be used to generate transgenic or knockout
 CC animals useful in the development and screening of therapeutically useful
 CC reagents. The PRO polypeptides and encoding nucleic acids can be used as
 CC molecular weight markers for protein electrophoresis, chromosome
 CC identification and tissue typing. The PRO polypeptides are useful to
 CC induce angiogenesis e.g wound healing, in the treatment of sports-related
 CC joint problems, articular cartilage defects, osteoarthritis or rheumatoid
 CC arthritis; diabetes; hyperinsulinaemia and hypoinsulinaemia. The
 CC antibodies may be used in various diagnostic, competitive binding and/or
 CC immunoprecipitation assays. The present sequence represents the amino
 CC acid sequence of a PRO polypeptide of the invention.

XX Query Match 2.0%; Score 8; DB 24; Length 596;
 XX Best Local Similarity 100.0%; Pred. No. 1.6e+02;

XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 81 SSGASTAT 88
 Db 39 SSGASTAT 46

```

RESULT 51
ABU72001
ID ABU72001 standard; Protein; 596 AA.
XX
AC ABU72001;
XX
DT 11-JUN-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO1342.
XX
KW Human; secreted and transmembrane polypeptide;
KW chromosome mapping; gene mapping; transgenic animal; knockout animal;
KW therapeutic agent screening; chromosome identification; tissue typing;
KW gene therapy.
XX
OS Homo sapiens.
XX
PN US2003018183-A1.
XX
PD 23-JAN-2003.
XX
PF 01-MAY-2002; 2002US-0063512.
XX
PR 06-DEC-2001; 2001US-0006867.
XX
PA (GETH ) GENENTECH INC.
XX
PI Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Watanabe CK, Wood WI;
XX
DR WPI; 2003-330984/31.
DR N-PSDB; ACA60412.
XX
XX
PT New secreted and transmembrane PRO polypeptides and nucleic acid
PT molecules encoding the polypeptides, useful in gene therapy or
PT preparing a medicament for treating a condition that is responsive to
PT the PRO polypeptide or antibody -
XX
PS Disclosure; Fig 100; 409pp; English.
XX
CC The invention describes novel isolated PRO polypeptides. The PRO
CC polypeptides or anti-PRO antibodies are useful in preparing a medicament
CC for treating a condition that is responsive to the PRO polypeptide or
CC antibody. The PRO nucleotide sequences may be used as hybridisation
CC probes in chromosome and gene mapping, or in generating antisense RNA
CC and DNA. PRO nucleic acids are also useful in preparing PRO polypeptides,
CC in assays to identify other proteins or molecules involved in binding
CC reaction, to generate transgenic animals or knockout animals, which in
CC turn are useful in the development and screening of therapeutically
CC useful reagents, for chromosome identification, and tissue typing. The
CC PRO polypeptides and nucleic acid molecules are also useful in gene
CC therapy, and as molecular weight markers for protein electrophoresis
CC purposes. The anti-PRO antibodies may be used in diagnostic assays for
CC PRO, or for the affinity purification of PRO from recombinant cell
CC culture or natural sources. This is the amino acid sequence of a novel
CC human secreted and transmembrane PRO polypeptide.
XX
SQ Sequence 596 AA;
Query Match 2.0%; Score 8; DB 24; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
Db |||||
39 SSGASTAT 46

RESULT 52
ABU72158
ID ABU72158 standard; Protein; 596 AA.
XX

```

```

AC ABU72158;
XX
DT 13-JUN-2003 (first entry)
XX
DE Human PRO polypeptide #50.
XX
KW Human; PRO polypeptide; secreted and transmembrane protein;
KW anti-PRO antibody; diagnostic assay; gene expression.
XX
OS Homo sapiens.
XX
PN US2003023042-A1.
XX
PD 30-JAN-2003.
XX
PF 01-MAY-2002; 2002US-0063502.
XX
PR 06-DEC-2001; 2001US-0006867.
XX
PA (GETH ) GENENTECH INC.
XX
PI Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Watanabe CK, Wood WI;
XX
DR WPI; 2003-331484/31.
DR N-PSDB; ACA63422.
XX
PT Novel monoclonal antibody that binds to secreted and transmembrane
PT polypeptide, useful for detecting and purifying the polypeptide and
PT also for treating conditions responsive to the antibody -
XX
PS Disclosure; Fig 100; 408pp; English.
XX
CC The present invention relates to the isolation of novel human PRO
CC polypeptides and the polynucleotide sequences encoding them. The
CC PRO polypeptides are secreted and transmembrane proteins. The PRO
CC polypeptides and polynucleotides are useful for preparing a
CC medicament useful in the treatment of a condition responsive to
CC anti-PRO antibody. Anti-PRO antibodies are useful in diagnostic
CC assays for PRO, by detecting its expression in specific cells,
CC tissues or serum, and for affinity purification of PRO from
CC recombinant cell culture or natural sources. ABU72109-ABU72192
CC represent the human PRO polypeptides of the invention.
XX
SQ Sequence 596 AA;
Query Match 2.0%; Score 8; DB 24; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
Db |||||
39 SSGASTAT 46

RESULT 53
ABU65723
ID ABU65723 standard; Protein; 596 AA.
XX
AC ABU65723;
XX
DT 19-MAY-2003 (first entry)
XX
DE Human secreted/transmembrane protein, SEQ ID 310.
XX
KW Human; PRO; secreted protein; transmembrane protein;
KW cytostatic; antiarthritic; osteopathic; adrenal tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW cervical tumour; liver tumour; TNF-alpha release; arthritis;
KW tumour necrosis factor alpha; chondrocyte cell; bone disorder;
KW cartilage disorder; sports injury.
XX
OS Homo sapiens.

```

XX	US2003036156-A1.		
PN		PR 01-APR-1998;	98US-080333P.
XX		PR 08-APR-1998;	98US-081049P.
PD		PR 08-APR-1998;	98US-081070P.
XX	20-FEB-2003.	PR 09-APR-1998;	98US-081195P.
PF		PR 15-APR-1998;	98US-081838P.
XX		PR 21-APR-1998;	98US-082568P.
XX	02-JUL-2002; 2002US-0188767.	PR 21-APR-1998;	98US-082569P.
XX		PR 22-APR-1998;	98US-082704P.
PR	16-SEP-1998;	PR 22-APR-1998;	98US-082797P.
PR	07-OCT-1998;	PR 28-APR-1998;	98US-083322P.
PR	01-DEC-1998;	PR 29-APR-1998;	98US-083495P.
PR	08-MAR-1999;	PR 29-APR-1998;	98US-083496P.
PR	14-MAY-1999;	PR 29-APR-1998;	98US-083499P.
PR	02-JUN-1999;	PR 29-APR-1998;	98US-083559P.
PR	01-SEP-1999;	PR 05-MAY-1998;	98US-084366P.
PR	15-SEP-1999;	PR 06-MAY-1998;	98US-084414P.
PR	01-DEC-1999;	PR 07-MAY-1998;	98US-084639P.
PR	02-DEC-1999;	PR 07-MAY-1998;	98US-084640P.
PR	30-DEC-1999;	PR 07-MAY-1998;	98US-084643P.
PR	05-JAN-2000;	PR 15-MAY-1998;	98US-085579P.
PR	18-FEB-2000;	PR 15-MAY-1998;	98US-085580P.
PR	18-FEB-2000;	PR 15-MAY-1998;	98US-085582P.
PR	22-FEB-2000;	PR 15-MAY-1998;	98US-085700P.
PR	24-FEB-2000;	PR 18-MAY-1998;	98US-086023P.
PR	01-MAR-2000;	PR 22-MAY-1998;	98US-086392P.
PR	02-MAR-2000;	PR 22-MAY-1998;	98US-086486P.
PR	15-MAR-2000;	PR 28-MAY-1998;	98US-087098P.
PR	30-MAR-2000;	PR 28-MAY-1998;	98US-087208P.
PR	17-MAY-2000;	PR 02-JUN-1998;	98US-087609P.
PR	22-MAY-2000;	PR 02-JUN-1998;	98US-087759P.
PR	30-MAY-2000;	PR 03-JUN-1998;	98US-087827P.
PR	02-JUN-2000;	PR 04-JUN-1998;	98US-088025P.
PR	28-JUL-2000;	PR 04-JUN-1998;	98US-088028P.
PR	24-AUG-2000;	PR 04-JUN-1998;	98US-088029P.
PR	08-NOV-2000;	PR 04-JUN-1998;	98US-088033P.
PR	01-DEC-2000;	PR 04-JUN-1998;	98US-088326P.
PR	20-DEC-2000;	PR 05-JUN-1998;	98US-088167P.
PR	28-FEB-2001;	PR 05-JUN-1998;	98US-088202P.
PR	01-JUN-2001;	PR 05-JUN-1998;	98US-088212P.
PR	20-JUN-2001;	PR 05-JUN-1998;	98US-088217P.
PR	29-JUN-2001;	PR 09-JUN-1998;	98US-088655P.
PR	09-JUL-2001;	PR 10-JUN-1998;	98US-088722P.
PR	29-AUG-2001;	PR 10-JUN-1998;	98US-088738P.
PR	18-SEP-1997;	PR 10-JUN-1998;	98US-088740P.
PR	18-SEP-1997;	PR 10-JUN-1998;	98US-088811P.
PR	17-OCT-1997;	PR 10-JUN-1998;	98US-088824P.
PR	21-OCT-1997;	PR 10-JUN-1998;	98US-088825P.
PR	24-OCT-1997;	PR 10-JUN-1998;	98US-088826P.
PR	24-OCT-1997;	PR 11-JUN-1998;	98US-088861P.
PR	28-OCT-1997;	PR 11-JUN-1998;	98US-088863P.
PR	28-OCT-1997;	PR 11-JUN-1998;	98US-088876P.
PR	28-OCT-1997;	PR 12-JUN-1998;	98US-089090P.
PR	29-OCT-1997;	PR 12-JUN-1998;	98US-089105P.
PR	31-OCT-1997;	PR 16-JUN-1998;	98US-089512P.
PR	31-OCT-1997;	PR 16-JUN-1998;	98US-089514P.
PR	13-NOV-1997;	PR 17-JUN-1998;	98US-089538P.
PR	21-NOV-1997;	PR 17-JUN-1998;	98US-089598P.
PR	24-NOV-1997;	PR 17-JUN-1998;	98US-089653P.
PR	24-NOV-1997;	PR 18-JUN-1998;	98US-089508P.
PR	11-DEC-1997;	PR 19-JUN-1998;	98US-089952P.
PR	12-DEC-1997;	PR 22-JUN-1998;	98US-090246P.
PR	17-DEC-1997;	PR 22-JUN-1998;	98US-090252P.
PR	18-DEC-1997;	PR 22-JUN-1998;	98US-090254P.
PR	10-MAR-1998;	PR 24-JUN-1998;	98US-090429P.
PR	11-MAR-1998;	PR 24-JUN-1998;	98US-090435P.
PR	11-MAR-1998;	PR 24-JUN-1998;	98US-090444P.
PR	20-MAR-1998;	PR 24-JUN-1998;	98US-090461P.
PR	20-MAR-1998;	PR 24-JUN-1998;	98US-090535P.
PR	27-MAR-1998;	PR 24-JUN-1998;	98US-090540P.
PR	27-MAR-1998;	PR 25-JUN-1998;	98US-090676P.
PR	31-MAR-1998;	PR 25-JUN-1998;	98US-090678P.
PR	31-MAR-1998;	PR 25-JUN-1998;	98US-090688P.
PR	01-APR-1998;	PR 25-JUN-1998;	98US-090690P.

PR 25-JUN-1998; 98US-090594P.
 PR 25-JUN-1998; 98US-090595P.
 PR 25-JUN-1998; 98US-090596P.
 PR 26-JUN-1998; 98US-090862P.
 PR 26-JUN-1998; 98US-090863P.
 PR 26-JUN-1998; 98US-091010P.
 PR 01-JUL-1998; 98US-091159P.
 PR 01-JUL-1998; 98US-0911544P.
 PR 02-JUL-1998; 98US-0911478P.
 PR 02-JUL-1998; 98US-091486P.
 PR 02-JUL-1998; 98US-091626P.
 PR 02-JUL-1998; 98US-091628P.
 PR 02-JUL-1998; 98US-091632P.
 PR 24-JUL-1998; 98US-094006P.
 PR 04-AUG-1998; 98US-095282P.
 PR 10-AUG-1998; 98US-095398P.
 PR 10-AUG-1998; 98US-096012P.
 PR 17-AUG-1998; 98US-096757P.
 PR 17-AUG-1998; 98US-096766P.
 PR 17-AUG-1998; 98US-096867P.
 PR 17-AUG-1998; 98US-096891P.
 PR 17-AUG-1998; 98US-096897P.
 PR 18-AUG-1998; 98US-096949P.
 PR 18-AUG-1998; 98US-096959P.
 PR 18-AUG-1998; 98US-097022P.
 PR 26-AUG-1998; 98US-097952P.
 PR 26-AUG-1998; 98US-097954P.
 PR 26-AUG-1998; 98US-097955P.
 PR 26-AUG-1998; 98US-097971P.
 PR 26-AUG-1998; 98US-097974P.
 PR 26-AUG-1998; 98US-098014P.
 PR 01-SEP-1998; 98US-098716P.
 PR 01-SEP-1998; 98US-098723P.
 PR 02-SEP-1998; 98US-098803P.
 PR 02-SEP-1998; 98US-098821P.
 PR 02-SEP-1998; 98US-098843P.
 PR 09-SEP-1998; 98US-099602P.
 PR 10-SEP-1998; 98US-099741P.

Query Match 2.0%; Score 8; DB 24; Length 596;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 SSGASTAT 88
 Db 39 SSGASTAT 46

RESULT 54
 ABU66056
 ID ABU66056 standard; Protein; 596 AA.

XX AC ABU66056;

XX DT 20-MAY-2003 (first entry)

DE XX Novel human secreted and transmembrane protein PRO1342.

XX KW Human; secreted protein; transmembrane protein; cytostatic;
 KW gene therapy; TNF-Agonist-Alpha; chondrocyte stimulator; tumour;
 KW adrenal tumour; lung tumour; colon tumour; breast tumour;
 KW prostate tumour; rectal tumour; cervical tumour; liver tumour.

OS XX Homo sapiens.

XX PN US2003036157-A1.

XX PD 20-FEB-2003.

PF 02-JUL-2002; 2002US-0188769.

XX 16-SEP-1998; 98WO-US19330.

PR 07-OCT-1998; 98WO-US21141.

PR 01-DEC-1998; 98WO-US25108.
 PR 08-MAR-1999; 99WO-US05028.
 PR 14-MAY-1999; 99WO-US10733.
 PR 02-JUN-1999; 99WO-US12252.
 PR 15-SEP-1999; 99WO-US20111.
 PR 15-SEP-1999; 99WO-US21090.
 PR 01-DEC-1999; 99WO-US28301.
 PR 02-DEC-1999; 99WO-US28551.
 PR 30-DEC-1999; 99WO-US31274.
 PR 05-JAN-2000; 2000WO-US00219.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 18-FEB-2000; 2000WO-US04342.
 PR 22-FEB-2000; 2000WO-US04414.
 PR 24-FEB-2000; 2000WO-US05004.
 PR 01-MAR-2000; 2000WO-US05601.
 PR 02-MAR-2000; 2000WO-US05841.
 PR 15-MAR-2000; 2000WO-US06884.
 PR 30-MAR-2000; 2000WO-US08439.
 PR 17-MAY-2000; 2000WO-US13705.
 PR 22-MAY-2000; 2000WO-US14042.
 PR 30-MAY-2000; 2000WO-US14941.
 PR 02-JUN-2000; 2000WO-US15264.
 PR 28-JUL-2000; 2000WO-US20710.
 PR 24-AUG-2000; 2000WO-US23328.
 PR 08-NOV-2000; 2000WO-US30952.
 PR 01-DEC-2000; 2000WO-US32678.
 PR 20-DEC-2000; 2000WO-US34956.
 PR 28-FEB-2001; 2001WO-US06520.
 PR 01-JUN-2001; 2001WO-US17800.
 PR 20-JUN-2001; 2001WO-US19692.
 PR 29-JUN-2001; 2001WO-US21066.
 PR 09-JUL-2001; 2001WO-US21735.
 PR 29-AUG-2001; 2001WO-US27099.
 PR 18-SEP-1997; 97US-059263P.
 PR 18-SEP-1997; 97US-059266P.
 PR 17-OCT-1997; 97US-062250P.
 PR 21-OCT-1997; 97US-063486P.
 PR 24-OCT-1997; 97US-063120P.
 PR 24-OCT-1997; 97US-063121P.
 PR 28-OCT-1997; 97US-063540P.
 PR 28-OCT-1997; 97US-063541P.
 PR 28-OCT-1997; 97US-063564P.
 PR 29-OCT-1997; 97US-063734P.
 PR 31-OCT-1997; 97US-063870P.
 PR 31-OCT-1997; 97US-064103P.
 PR 13-NOV-1997; 97US-065311P.
 PR 21-NOV-1997; 97US-066120P.
 PR 24-NOV-1997; 97US-066466P.
 PR 11-DEC-1997; 97US-066772P.
 PR 12-DEC-1997; 97US-069335P.
 PR 17-DEC-1997; 97US-069425P.
 PR 18-DEC-1997; 97US-069870P.
 PR 10-MAR-1998; 98US-068017P.
 PR 11-MAR-1998; 98US-077632P.
 PR 11-MAR-1998; 98US-077649P.
 PR 20-MAR-1998; 98US-078886P.
 PR 27-MAR-1998; 98US-078939P.
 PR 27-MAR-1998; 98US-079664P.
 PR 31-MAR-1998; 98US-079786P.
 PR 31-MAR-1998; 98US-080107P.
 PR 31-MAR-1998; 98US-080194P.
 PR 01-APR-1998; 98US-080327P.
 PR 01-APR-1998; 98US-080333P.
 PR 08-APR-1998; 98US-081049P.
 PR 08-APR-1998; 98US-081070P.
 PR 09-APR-1998; 98US-081195P.
 PR 15-APR-1998; 98US-081838P.
 PR 21-APR-1998; 98US-082568P.
 PR 21-APR-1998; 98US-082569P.
 PR 22-APR-1998; 98US-082704P.
 PR 22-APR-1998; 98US-082797P.

PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083599P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087659P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.

PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098943P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.
PR 10-SEP-1998; 98US-099754P.
PR 10-SEP-1998; 98US-099763P.

Query Match 2.0%; Score 8; DB 24; Length 596;

Best Local Similarity 100.0%; Pred. No. 1.6e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
Db 39 SSGASTAT 46

RESULT 55

ABU67560

ID ABU67560 standard; Protein; 596 AA.

XX AC ABU67560;

XX DT 29-MAY-2003 (first entry)

XX DE Human secreted/transmembrane protein (PRO) #155.

XX KW Human; secreted and transmembrane protein; PRO; TNF-alpha;
tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
tissue typing.

XX OS Homo sapiens.

XX PN US2003036162-A1.

XX PD 20-FEB-2003.

XX PF 12-JUL-2002; 2002US-0194423.

XX PR 16-SEP-1998; 98WO-US19330.

XX PR 07-OCT-1998; 98WO-US21141.

XX PR 01-DEC-1998; 98WO-US25108.

XX PR 08-MAR-1999; 99WO-US05028.

XX PR 14-MAY-1999; 99WO-US10733.

XX PR 02-JUN-1999; 99WO-US12252.

XX PR 01-SEP-1999; 99WO-US20111.

XX PR 15-SEP-1999; 99WO-US21090.

XX PR 01-DEC-1999; 99WO-US28301.

XX PR 02-DEC-1999; 99WO-US28551.

PR 30-DEC-1999; 99WO-US31274.
 PR 05-JAN-2000; 2000WO-US00219.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 18-FEB-2000; 2000WO-US04342.
 PR 22-FEB-2000; 2000WO-US04414.
 PR 24-FEB-2000; 2000WO-US05004.
 PR 01-MAR-2000; 2000WO-US05601.
 PR 01-MAR-2000; 2000WO-US05841.
 PR 15-MAR-2000; 2000WO-US06884.
 PR 30-MAR-2000; 2000WO-US08439.
 PR 17-MAY-2000; 2000WO-US13705.
 PR 22-MAY-2000; 2000WO-US14042.
 PR 30-MAY-2000; 2000WO-US14941.
 PR 02-JUN-2000; 2000WO-US15264.
 PR 28-JUL-2000; 2000WO-US20710.
 PR 24-AUG-2000; 2000WO-US23328.
 PR 08-NOV-2000; 2000WO-US30952.
 PR 01-DEC-2000; 2000WO-US32678.
 PR 28-DEC-2000; 2000WO-US34956.
 PR 28-DEC-2000; 2001WO-US06520.
 PR 01-JUN-2001; 2001WO-US17800.
 PR 20-JUN-2001; 2001WO-US19692.
 PR 29-JUN-2001; 2001WO-US21066.
 PR 09-JUL-2001; 2001WO-US21735.
 PR 28-AUG-2001; 2001WO-US27099.
 PR 26-JUN-1998; 98US-0105413.
 PR 07-OCT-1998; 98US-0168978.
 PR 06-NOV-1998; 98US-0187368.
 PR 07-DEC-1998; 98US-0202054.
 PR 03-MAR-1999; 98US-0254311.
 PR 14-MAY-1999; 98US-0311832.
 PR 14-MAY-1999; 99US-0380137.
 PR 25-AUG-1999; 99US-0380138.
 PR 25-AUG-1999; 99US-0380139.
 PR 25-AUG-1999; 99US-0380142.
 PR 18-OCT-1999; 99US-0403297.
 PR 12-NOV-1999; 99US-0423844.
 PR 22-AUG-2000; 2000US-0644848.
 PR 18-SEP-2000; 2000US-0664610.
 PR 18-SEP-2000; 2000US-0665350.
 PR 08-NOV-2000; 2000US-0709238.
 PR 20-DEC-2000; 2000US-0747259.
 PR 22-MAR-2001; 2001US-0816744.
 PR 10-MAY-2001; 2001US-0854208.
 PR 10-MAY-2001; 2001US-0854280.
 PR 25-MAY-2001; 2001US-0866028.
 PR 05-JUN-2001; 2001US-0874503.
 PR 18-JUL-2001; 2001US-0908827.
 PR 30-JUL-2001; 2001US-0918585.
 PR 06-AUG-2001; 2001US-0924419.
 PR 13-AUG-2001; 2001US-0929404.
 PR 16-AUG-2001; 2001US-0931836.
 PR 28-AUG-2001; 2001US-0941992.
 PR 04-SEP-2001; 2001US-0946374.
 PR 15-JAN-2002; 2002US-0052586.

(GETH) GENENTECH INC.

Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
 Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;

WPI; 2003-332039/31.

N-PSDB; ACA05854.

New secreted and transmembrane PRO polypeptides and nucleic acids,
 useful in gene therapy, in chromosome and gene mapping, as chromosome
 markers, in tissue typing, and in chromosome identification -

Claim 11; Fig 310; 706pp; English.

The invention discloses human nucleic acids encoding secreted and
 transmembrane (PRO) polypeptides. Also disclosed is an antibody that
 specifically binds to the PRO polypeptide, a method for stimulating the

CC release of tumour necrosis factor alpha (TNF-alpha) from human blood by
 CC contacting the blood a PRO polypeptide, a method for stimulating the
 CC proliferation or differentiation of chondrocyte cells by contacting the
 CC cells with a PRO polypeptide, a method for detecting the presence of a
 CC tumour in a mammal and an oligonucleotide probe derived from any of the
 CC PRO nucleotide sequences. The nucleotide sequences are useful as probes,
 CC in chromosome and gene mapping, in generating antisense RNA and DNA, in
 CC preparing PRO polypeptides by recombinant techniques and in gene therapy
 CC (e.g. for replacement of defective gene). The PRO polypeptides are useful
 CC as molecular weight markers for protein electrophoresis purposes, for
 CC chromosome identification, as chromosome markers, as therapeutic agents,
 CC for stimulating the release of TNF-alpha from human blood, for
 CC stimulating the proliferation or differentiation of chondrocytes and
 CC detecting the presence of a tumour. The PRO polypeptides and nucleic
 CC acids may also be used diagnostically for tissue typing. The sequences
 CC presented in ABU67406-ABU67710 are the PRO polypeptides of the invention.

XX Sequence 596 AA;

Query Match 2.0%; Score 8; DB 24; Length 596;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 SSGASTAT 88
 Db 39 SSGASTAT 46
 |||||

RESULT 56

ABU65418
 ID ABU65418 standard; Protein; 596 AA.

XX AC ABU65418;

XX DT 16-MAY-2003 (first entry)

XX DE Human PRO polypeptide #155.

XX KW Human; PRO; cytostatic; chromosome mapping; gene mapping;
 KW protein electrophoresis; tumour necrosis factor-alpha; TNF-alpha; blood;
 KW chondrocyte differentiation; chondrocyte proliferation; tumour.

XX OS Homo sapiens.

XX FN US2003032102-A1.

XX PD 13-FEB-2003.

XX PF 17-JUN-2002; 2002US-0173697.

XX PR 16-SEP-1998; 98WO-US19330.

PR 07-OCT-1998; 98WO-US21141.

PR 01-DEC-1998; 98WO-US25108.

PR 08-MAR-1999; 99WO-US05028.

PR 14-MAY-1999; 99WO-US10733.

PR 02-JUN-1999; 99WO-US12252.

PR 01-SEP-1999; 99WO-US20111.

PR 15-SEP-1999; 99WO-US21090.

PR 01-DEC-1999; 99WO-US28301.

PR 02-DEC-1999; 99WO-US28551.

PR 30-DEC-1999; 99WO-US31274.

PR 05-JAN-2000; 2000WO-US00219.

PR 18-FEB-2000; 2000WO-US04341.

PR 22-FEB-2000; 2000WO-US04414.

PR 24-FEB-2000; 2000WO-US05004.

PR 01-MAR-2000; 2000WO-US05601.

PR 02-MAR-2000; 2000WO-US05841.

PR 15-MAR-2000; 2000WO-US06884.

PR 30-MAR-2000; 2000WO-US08439.

PR 17-MAY-2000; 2000WO-US13705.

PR 22-MAY-2000; 2000WO-US14042.

PR 30-MAY-2000; 2000WO-US14941.

PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30352.
PR 01-DEC-2000; 2000WO-US32578.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 08-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 28-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 08-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 15-APR-1998; 98US-081159P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082569P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088555P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 26-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.

PR 26-AUG-1998; 98US-098014P.
 PR 01-SEP-1998; 98US-098716P.
 PR 01-SEP-1998; 98US-098723P.
 PR 02-SEP-1998; 98US-098803P.
 PR 02-SEP-1998; 98US-098821P.
 PR 02-SEP-1998; 98US-098843P.
 PR 09-SEP-1998; 98US-099022P.
 PR 10-SEP-1998; 98US-099741P.
 PR 10-SEP-1998; 98US-099754P.
 PR 10-SEP-1998; 98US-099763P.
 PR 10-SEP-1998; 98US-099812P.

Query Match 2.0%; Score 8; DB 24; Length 596;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88

Db 39 SSGASTAT 46

RESULT 57

ABUS8554

ID ABUS8554 standard; Protein; 596 AA.

XX AC ABUS8554;

XX DT 15-APR-2003 (first entry)

XX DE Human PRO polypeptide #155.

XX KW Human; PRO; cytostatic; tumour; cancer; breast; lung; stomach;

XX KW liver; dog; cat; cow; horse; sheep; pig; goat; rabbit; ADEPT;

XX KW antibody-dependent enzyme mediated prodrug therapy.

XX OS Homo sapiens.

XX PN US2003027272-A1.

XX PD 06-FEB-2003.

XX PF 21-JUN-2002; 2002US-0176492.

XX PR 16-SEP-1998; 98WO-US19330.

PR 07-OCT-1998; 98WO-US21141.

PR 01-DEC-1998; 98WO-US25108.

PR 08-MAR-1999; 98WO-US05028.

PR 10-MAR-1999; 98WO-US05190.

PR 14-MAY-1999; 98WO-US10733.

PR 02-JUN-1999; 98WO-US12252.

PR 01-SEP-1999; 98WO-US20111.

PR 15-SEP-1999; 98WO-US21090.

PR 30-NOV-1999; 98WO-US28313.

PR 01-DEC-1999; 98WO-US28301.

PR 02-DEC-1999; 98WO-US28551.

PR 05-JAN-2000; 98WO-US31274.

PR 18-FEB-2000; 2000WO-US04341.

PR 22-FEB-2000; 2000WO-US04342.

PR 24-FEB-2000; 2000WO-US04414.

PR 01-MAR-2000; 2000WO-US05004.

PR 02-MAR-2000; 2000WO-US05601.

PR 10-MAR-2000; 2000WO-US05841.

PR 15-MAR-2000; 2000WO-US06319.

PR 30-MAR-2000; 2000WO-US06884.

PR 17-MAY-2000; 2000WO-US08439.

PR 22-MAY-2000; 2000WO-US13705.

PR 30-MAY-2000; 2000WO-US14042.

PR 02-JUN-2000; 2000WO-US14941.

PR 28-JUL-2000; 2000WO-US15264.

PR 24-AUG-2000; 2000WO-US20710.

PR 08-NOV-2000; 2000WO-US23328.

PR 08-NOV-2000; 2000WO-US30952.

PR 10-NOV-2000; 2000WO-US30873.
 PR 01-DEC-2000; 2000WO-US32678.
 PR 20-DEC-2000; 2000WO-US34956.
 PR 28-FEB-2001; 2001WO-US06520.
 PR 01-JUN-2001; 2001WO-US17800.
 PR 20-JUN-2001; 2001WO-US19692.
 PR 29-JUN-2001; 2001WO-US21066.
 PR 09-JUL-2001; 2001WO-US21735.
 PR 18-SEP-1997; 97US-059263P.
 PR 18-SEP-1997; 97US-059266P.
 PR 17-OCT-1997; 97US-062250P.
 PR 21-OCT-1997; 97US-063486P.
 PR 24-OCT-1997; 97US-063120P.
 PR 28-OCT-1997; 97US-063540P.
 PR 28-OCT-1997; 97US-063541P.
 PR 28-OCT-1997; 97US-063544P.
 PR 29-OCT-1997; 97US-063564P.
 PR 29-OCT-1997; 97US-063734P.
 PR 31-OCT-1997; 97US-063870P.
 PR 31-OCT-1997; 97US-064103P.
 PR 13-NOV-1997; 97US-065311P.
 PR 21-NOV-1997; 97US-066120P.
 PR 24-NOV-1997; 97US-066466P.
 PR 24-NOV-1997; 97US-066772P.
 PR 11-DEC-1997; 97US-069335P.
 PR 12-DEC-1997; 97US-069425P.
 PR 17-DEC-1997; 97US-069870P.
 PR 18-DEC-1997; 97US-068017P.
 PR 10-MAR-1998; 98US-077450P.
 PR 11-MAR-1998; 98US-077632P.
 PR 11-MAR-1998; 98US-077649P.
 PR 20-MAR-1998; 98US-078886P.
 PR 20-MAR-1998; 98US-078939P.
 PR 27-MAR-1998; 98US-079664P.
 PR 27-MAR-1998; 98US-079786P.
 PR 31-MAR-1998; 98US-080194P.
 PR 01-APR-1998; 98US-080327P.
 PR 01-APR-1998; 98US-080333P.
 PR 08-APR-1998; 98US-081049P.
 PR 08-APR-1998; 98US-081070P.
 PR 09-APR-1998; 98US-081195P.
 PR 15-APR-1998; 98US-081838P.
 PR 21-APR-1998; 98US-082568P.
 PR 21-APR-1998; 98US-082569P.
 PR 22-APR-1998; 98US-082704P.
 PR 22-APR-1998; 98US-082797P.
 PR 29-APR-1998; 98US-083322P.
 PR 29-APR-1998; 98US-083495P.
 PR 29-APR-1998; 98US-083496P.
 PR 29-APR-1998; 98US-083499P.
 PR 29-APR-1998; 98US-083559P.
 PR 06-MAY-1998; 98US-084366P.
 PR 07-MAY-1998; 98US-084639P.
 PR 07-MAY-1998; 98US-084640P.
 PR 07-MAY-1998; 98US-084643P.
 PR 15-MAY-1998; 98US-085579P.
 PR 15-MAY-1998; 98US-085580P.
 PR 15-MAY-1998; 98US-085582P.
 PR 15-MAY-1998; 98US-085700P.
 PR 18-MAY-1998; 98US-086023P.
 PR 22-MAY-1998; 98US-086392P.
 PR 22-MAY-1998; 98US-086486P.
 PR 28-MAY-1998; 98US-087098P.
 PR 28-MAY-1998; 98US-087208P.
 PR 02-JUN-1998; 98US-087609P.
 PR 02-JUN-1998; 98US-087753P.
 PR 03-JUN-1998; 98US-087827P.
 PR 04-JUN-1998; 98US-088025P.
 PR 04-JUN-1998; 98US-088028P.
 PR 04-JUN-1998; 98US-088029P.

```
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088157P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 10-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.
Query Match 2.0%; Score 8; DB 24; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 81 SSGASTAT 88
DB 39 SSGASTAT 46
RESULT 58
ABUS6090
ID ABUS6090 standard; Protein; 596 AA.
XX
AC ABUS6090;
XX
DT 26-MAR-2003 (first entry)
XX
DE Human secreted/transmembrane protein, PRO1342.
XX
Human; secreted protein; transmembrane protein; PRO;
KW antiarthritic; vulnary; tumour necrosis factor-alpha;
KW chondrocyte cell proliferation; chondrocyte cell differentiation;
KW tumour; adrenal tumour; lung tumour; colon tumour; breast tumour;
KW prostate tumour; rectal tumour; cervical tumour; liver tumour;
KW bone disorder; cartilage disorder; arthritis; sports injury.
XX
OS Homo sapiens.
XX
PN US2003022298-A1.
XX
PD 30-JAN-2003.
XX
PF 20-JUN-2002; 2002US-0176913.
XX
05-NOV-1997; 97WO-US20069.
PR 10-SEP-1998; 98WO-US18824.
PR 14-SEP-1998; 98WO-US19177.
PR 16-SEP-1998; 98WO-US19330.
PR 17-SEP-1998; 98WO-US19437.
PR 07-OCT-1998; 98WO-US21141.
PR 20-NOV-1998; 98WO-US24855.
PR 01-DEC-1998; 98WO-US25108.
PR 05-JAN-1999; 99WO-US00106.
PR 08-MAR-1999; 99WO-US05028.
PR 10-MAR-1999; 99WO-US05190.
PR 20-APR-1999; 99WO-US08615.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28409.
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 22-DEC-1999; 99WO-US30720.
PR 30-DEC-1999; 99WO-US31243.
PR 30-DEC-1999; 99WO-US31274.
```

PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 10-MAR-2000; 2000WO-US06319.
PR 15-MAR-2000; 2000WO-US06884.
PR 21-MAR-2000; 2000WO-US07532.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 02-JUN-2000; 2000WO-US14941.
PR 30-MAY-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 11-AUG-2000; 2000WO-US22031.
PR 23-AUG-2000; 2000WO-US23522.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 10-NOV-2000; 2000WO-US30873.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 22-MAR-2001; 2001WO-US09552.
PR 25-MAY-2001; 2001WO-US17092.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 1-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 20-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.

PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086192P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089598P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.

Query Match 2.0%; Score 8; DB 24; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches	8;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	81	SSGASTAT 88							
Db	39	SSGASTAT 46							
RESULT 59									
ABU57085									
ID	ABU57085 standard; Protein; 596 AA.								
XX									
AC	ABU57085;								
DT	04-APR-2003 (first entry)								
XX									
DE	Human PRO polypeptide #155.								
XX									
KW	Human; PRO; tumour necrosis factor-alpha; blood; cancer;								
KW	chondrocyte cell; tumour; adrenal tumour; lung; colon; breast; prostate;								
KW	kidney; rectum; cervix; liver; bone disorder; cartilage disorder;								
KW	arthritis; sports injury; genetic disorder; antiarthritic; vulnery.								
XX									
OS	Homo sapiens.								
XX									
PN	US2003027280-A1.								
XX									
PD	06-FEB-2003.								
XX									
PF	20-JUN-2002; 2002US-0176993.								
XX									
PR	16-SEP-1998;	98WO-US19330.							
PR	07-OCT-1998;	98WO-US21141.							
PR	01-DEC-1998;	98WO-US25108.							
PR	08-MAR-1999;	99WO-US05028.							
PR	14-MAY-1999;	99WO-US10733.							
PR	02-JUN-1999;	99WO-US12252.							
PR	01-SEP-1999;	99WO-US20111.							
PR	15-SEP-1999;	99WO-US21090.							
PR	01-DEC-1999;	99WO-US28301.							
PR	02-DEC-1999;	99WO-US28551.							
PR	30-DEC-1999;	99WO-US31274.							
PR	05-JAN-2000;	2000WO-US00219.							
PR	18-FEB-2000;	2000WO-US04341.							
PR	18-FEB-2000;	2000WO-US04342.							
PR	22-FEB-2000;	2000WO-US04414.							
PR	24-FEB-2000;	2000WO-US05004.							
PR	01-MAR-2000;	2000WO-US05601.							
PR	02-MAR-2000;	2000WO-US05841.							
PR	15-MAR-2000;	2000WO-US06884.							
PR	30-MAR-2000;	2000WO-US08439.							
PR	17-MAY-2000;	2000WO-US13705.							
PR	22-MAY-2000;	2000WO-US14042.							
PR	30-MAY-2000;	2000WO-US14941.							
PR	02-JUN-2000;	2000WO-US15264.							
PR	28-JUL-2000;	2000WO-US20710.							
PR	24-AUG-2000;	2000WO-US23328.							
PR	08-NOV-2000;	2000WO-US23952.							
PR	01-DEC-2000;	2000WO-US2678.							
PR	20-DEC-2000;	2000WO-US34956.							
PR	28-FEB-2001;	2001WO-US06520.							
PR	01-JUN-2001;	2001WO-US17800.							
PR	20-JUN-2001;	2001WO-US19692.							
PR	29-JUN-2001;	2001WO-US21066.							
PR	09-JUL-2001;	2001WO-US21735.							
PR	29-AUG-2001;	2001WO-US27099.							
PR	18-SEP-1997;	97US-059263P.							
PR	18-SEP-1997;	97US-059266P.							
PR	17-OCT-1997;	97US-062250P.							
PR	21-OCT-1997;	97US-063486P.							
PR	24-OCT-1997;	97US-063120P.							
PR	24-OCT-1997;	97US-063121P.							
PR	28-OCT-1997;	97US-063540P.							
PR	28-OCT-1997;	97US-063541P.							

PR	28-OCT-1997;	97US-063544P.
PR	28-OCT-1997;	97US-063564P.
PR	29-OCT-1997;	97US-063734P.
PR	31-OCT-1997;	97US-063870P.
PR	31-OCT-1997;	97US-064103P.
PR	13-NOV-1997;	97US-065311P.
PR	21-NOV-1997;	97US-066120P.
PR	24-NOV-1997;	97US-066466P.
PR	24-NOV-1997;	97US-066772P.
PR	11-DEC-1997;	97US-069335P.
PR	12-DEC-1997;	97US-069425P.
PR	17-DEC-1997;	97US-069870P.
PR	18-DEC-1997;	97US-068017P.
PR	10-MAR-1998;	98US-077450P.
PR	11-MAR-1998;	98US-077632P.
PR	11-MAR-1998;	98US-077649P.
PR	20-MAR-1998;	98US-078866P.
PR	20-MAR-1998;	98US-078939P.
PR	27-MAR-1998;	98US-079664P.
PR	27-MAR-1998;	98US-079786P.
PR	31-MAR-1998;	98US-080107P.
PR	31-MAR-1998;	98US-080194P.
PR	01-APR-1998;	98US-080327P.
PR	01-APR-1998;	98US-080333P.
PR	08-APR-1998;	98US-081049P.
PR	08-APR-1998;	98US-081070P.
PR	09-APR-1998;	98US-081195P.
PR	15-APR-1998;	98US-081838P.
PR	21-APR-1998;	98US-082568P.
PR	21-APR-1998;	98US-082569P.
PR	22-APR-1998;	98US-082704P.
PR	28-APR-1998;	98US-082797P.
PR	28-APR-1998;	98US-083322P.
PR	29-APR-1998;	98US-083495P.
PR	29-APR-1998;	98US-083496P.
PR	29-APR-1998;	98US-083499P.
PR	05-MAY-1998;	98US-083559P.
PR	06-MAY-1998;	98US-084366P.
PR	07-MAY-1998;	98US-084414P.
PR	07-MAY-1998;	98US-084639P.
PR	07-MAY-1998;	98US-084640P.
PR	07-MAY-1998;	98US-084643P.
PR	15-MAY-1998;	98US-085579P.
PR	15-MAY-1998;	98US-085580P.
PR	15-MAY-1998;	98US-085582P.
PR	18-MAY-1998;	98US-086023P.
PR	22-MAY-1998;	98US-086392P.
PR	22-MAY-1998;	98US-086486P.
PR	28-MAY-1998;	98US-087098P.
PR	28-MAY-1998;	98US-087208P.
PR	02-JUN-1998;	98US-087609P.
PR	02-JUN-1998;	98US-087759P.
PR	03-JUN-1998;	98US-087827P.
PR	04-JUN-1998;	98US-088025P.
PR	04-JUN-1998;	98US-088028P.
PR	04-JUN-1998;	98US-088029P.
PR	04-JUN-1998;	98US-088033P.
PR	04-JUN-1998;	98US-088326P.
PR	05-JUN-1998;	98US-088167P.
PR	05-JUN-1998;	98US-088202P.
PR	05-JUN-1998;	98US-088212P.
PR	05-JUN-1998;	98US-088217P.
PR	09-JUN-1998;	98US-088655P.
PR	10-JUN-1998;	98US-088722P.
PR	10-JUN-1998;	98US-088738P.
PR	10-JUN-1998;	98US-088740P.
PR	10-JUN-1998;	98US-088811P.
PR	10-JUN-1998;	98US-088824P.
PR	10-JUN-1998;	98US-088825P.
PR	10-JUN-1998;	98US-088826P.
PR	11-JUN-1998;	98US-088861P.
PR	11-JUN-1998;	98US-088863P.

PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090443P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091526P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 02-JUL-1998; 98US-091633P.
PR 04-AUG-1998; 98US-094006P.
PR 10-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.
PR 10-SEP-1998; 98US-099754P.
PR 10-SEP-1998; 98US-099763P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090443P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091526P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 02-JUL-1998; 98US-091633P.
PR 04-AUG-1998; 98US-094006P.
PR 10-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.
PR 10-SEP-1998; 98US-099754P.
PR 10-SEP-1998; 98US-099763P.

Query Match
Best Local Similarity 2.0%; Score 8; DB 24; Length 596;
Matches 8; Conservativity 100.0%; Pred. No. 1.6e+02;
Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
DB 39 SSGASTAT 46

RESULT 60
ABU10664
ID ABU10664 standard; Protein; 596 AA.
XX
AC ABU10664;
XX
DT 03-FEB-2003 (first entry)
XX
DE Human secreted/transmembrane protein #155.
XX
KW Human; secreted and transmembrane protein; blood;
KW tumour necrosis factor-alpha; chondrocyte cell proliferation;
KW chondrocyte cell differentiation; tumour; adrenal tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW cervical tumour; liver tumour; bone disorder; cartilage disorder;
KW arthritis; sports injury.
XX
OS Homo sapiens.
XX
FN US2002127584-A1.
XX
PD 12-SEP-2002.
XX
PF 15-JAN-2002; 2002US-0052586.
XX
PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 06-JAN-1999; 2000WO-US00219.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21030.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19632.
PR 29-JUN-2001; 2001WO-US21086.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.

PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 05-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083455P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
XX

(GETH) GENENTECH INC.

PI Baker KP, Chen J, Deenoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WJ, Zhang Z;

WPI; 2003-066893/06.

N-PSDB; ABX16740.

XX Novel isolated PRO polypeptides e.g., PRO1079, PRO827, PRO791, PRO1131,
PT PRO116, PRO1183, PRO1343, PRO1760, PRO1567 or PRO4333, useful for
PT stimulating release of tumor necrosis factor-alpha from human blood -

XX Claim 11; Fig 310; 701pp; English.

XX The invention relates to an isolated PRO polypeptide comprising at least
CC 80% sequence identity to the protein sequences appearing as ABU0510-
CC ABU0814 (including a version lacking its associated signal peptide, or
CC an isolated extracellular domain of a PRO polypeptide with or without
CC its associated signal peptide. Also included are the nucleic acids
CC encoding the PRO proteins (being secreted and transmembrane proteins)
CC appearing as ABX16586-ABX16590, PRO expression vectors, host cells,
CC chimeric PRO fusion proteins, an anti-PRO antibody and a PRO
CC derived oligonucleotide sequence. The PRO polypeptides are useful for
CC stimulating release of tumor necrosis factor-alpha from human blood.
CC The PRO polypeptide PRO6029 is useful for stimulating proliferation or
CC differentiation of chondrocyte cells. The PRO polypeptides as specified
CC in the specification and having differential expression in tumor cells,
CC are useful for detecting presence of tumor in a mammal (such as adrenal
CC tumour, lung tumour, colon tumour, breast tumour, prostate tumour, rectal
CC tumour, cervical tumour or liver tumour. The PRO polypeptide PRO6029 is
CC useful for treating various bone and/or cartilage disorders such as
CC arthritis, and sports injuries. The PRO polypeptides are useful for
CC screening compounds to identify ant/agonists. PRO nucleic acids
CC are useful as hybridisation probes, in chromosome and gene mapping,
CC in the generation of anti-sense RNA and DNA, for the preparation of PRO

CC polypeptides and for generating knock-out animals. The present
CC sequence represents a PRO polypeptide.

SQ Sequence 596 AA;

Query Match 2.0%; Score 8; DB 24; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
|||||

Db 39 SSGASTAT 46
|||||

RESULT 61

ABU05975
ID ABU05975 standard; Protein; 643 AA.

XX AC ABU05975;

XX DT 08-APR-2003 (first entry)

XX DE M. tuberculosis and M. leprae marker protein #626.

XX KW Mycobacterioses; survival; virulence; protective antigen; vaccine;
KW Mycobacterial disease; tuberculosis; leprosy.

XX OS Mycobacterium tuberculosis.

XX OS Mycobacterium leprae.

XX FN WO200274903-A2.

XX PD 26-SEP-2002.

XX PF 22-FEB-2002; 2002WO-IB01973.

XX PR 22-FEB-2001; 2001US-270123P.

XX PA (INSP) INST PASTEUR.

XX PI Cole S;

XX DR WPI; 2002-759885/82.

XX Identifying and selecting genes for survival or virulence of
PT mycobacteria by a comparative genomic analysis of the sequences of
PT Mycobacterium tuberculosis and M. leprae -

XX Claim 17; Page 844-845; 874pp; English.

XX This invention relates to a novel method for identifying essential genes
CC for survival or virulence of mycobacteria species. The method comprises
CC aligning the genomic sequence of a first mycobacterium species on a
CC genomic sequence of a second mycobacterium species and selecting a
CC polynucleotide sequence that is highly conserved in both genomes with no
CC counterparts in other bacterial sequences and that corresponds
CC to an essential gene for the survival or virulence of mycobacterium
CC species. The method of the invention is useful for detecting M.
CC tuberculosis or M. leprae infection. The method reduces the number of
CC potential new targets and protective antigens for new drugs and vaccine
CC compositions to treat and prevent mycobacterial diseases, particularly
CC tuberculosis and leprosy. The present sequence represents a marker
CC protein from Mycobacterium tuberculosis and Mycobacterium leprae
CC identified using the method of the invention.

SQ Sequence 643 AA;

Query Match 2.0%; Score 8; DB 23; Length 643;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 LGQLAAAV 348
|||||

Db 29 LGQLAAAV 36

RESULT 62

ID AAB46707 standard; Protein; 892 AA.

XX AC AAB46707;

XX DT 12-APR-2001 (first entry)

XX DE Herpes simplex virus DNA polymerase protein fragment SEQ ID NO 16.

XX KW Genome; thermophilic enzyme; washing powder; bleaching.

XX OS Herpesvirus.

XX PN WO200075335-A2.

XX PD 14-DEC-2000.

XX PF 02-JUN-2000; 2000WO-IB00893.

XX PR 02-JUN-1999; 99US-0137120.

XX PA (DECO-) DECODE GENETICS EHF.

XX PI Hjorleifsdottir S, Hreggvidsson GO, Fridjonsson OH, Aevaresson A;
XX PI Kristjansson JK;

XX DR WPI; 2001-061727/07.

XX PT Isolated nucleic acid from bacteriophage RM 378 encoding enzymes useful
XX PT in recombinant DNA technology

XX PS Disclosure; Figure 3A-P; 42pp; English.

XX CC This invention describes a novel isolated nucleic molecule (I) comprising
 CC the genome of bacteriophage RM 378. The invention also describes (1) an
 CC isolated nucleic acid which encodes a polypeptide obtainable from
 CC bacteriophage RM 378, or its active derivative or fragment; (2) an
 CC isolated bacteriophage RM 378 (II); (3) a DNA construct (III) comprising
 CC operatively linked to a regulatory sequence; (4) a host cell comprising
 CC (III); and (5) an isolated polypeptide (IV) obtainable from (II), its
 CC active derivative or fragment. Bacteriophage RM 378 is useful for
 CC producing thermophilic enzymes useful in DNA research and commercial
 CC settings (e.g. proteases and lipases used in washing powder, hydrolytic
 CC enzymes used in bleaching). The isolated nucleic acid molecules and
 CC vectors are useful in the manufacture of encoded polypeptide, as probes
 CC for isolating homologous sequences (e.g. from other bacteriophage
 CC species), as well as for detecting the presence of the bacteriophage in
 CC a culture of host cells. The polypeptides can be used as a molecular
 CC weight marker on SDS-PAGE gels or on molecular sieve gel filtration
 CC columns. Because the host organism of the RM378 bacteriophage is a
 CC thermophile, the enzymes and proteins of the RM378 bacteriophage are
 CC significantly more thermostable than those of other (e.g. mesophilic)
 CC bacteriophages, such as the T4 bacteriophage of *Escherichia coli*. The
 CC enhanced stability of the enzymes and proteins of RM378 bacteriophage
 CC allows their use under temperature conditions which would be prohibitive
 CC for other enzymes, thus increasing the range of conditions which can be
 CC employed not only in DNA research but also in commercial settings.

XX SQ Sequence 892 AA;

Query Match 2.0%; Score 8; DB 22; Length 892;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 TLAQPAAP 102

| | | | |

Db 38 TLAQPAAP 45

| | | | |

RESULT 63

ID AAW78898 standard; Protein; 898 AA.

XX AC AAW78898;

XX DT 25-MAR-2003 (updated)

XX DT 21-DEC-1998 (first entry)

XX DE Rat UNC-5 homologue UNC5H-1.

XX KW UNC-5; UNC5H-1; rat; netrin receptor; cell migration;
XX KW axon guidance; diagnosis; therapy.

XX OS Rattus sp.

XX FH Key Location/Qualifiers

XX FT Peptide 580..594

XX FT /note= "peptide used to raise rabbit
XX FT polyclonal antisera"

XX PN WO9837085-A1.

XX PD 27-AUG-1998.

XX PF 19-FEB-1998; 98WO-US03143.

XX PR 19-FEB-1997; 97US-0808982.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Tessier-lavigne M, Leonardo ED, Hinck L, Masu M, Keinomasu K;

XX DR WPI; 1998-495364/42.

XX DR N-PSDB; AAV52940.

XX PT Netrin-binding, vertebrate proteins - useful for diagnosis, therapy
XX PT and the biopharmaceutical industry

XX PS Claim 1; Page 19-22; 32pp; English.

XX CC UNC5H-1 and UNC5H-2 (see AAW78900) are rat homologues of
 CC Caenorhabditis elegans UNC-5 protein. Their amino acid sequences
 CC were deduced from isolated unc5h cDNA clones (see AAV52940 and
 CC AAV52942) isolated from an E18 brain cDNA library. The predicted
 CC proteins show similarity with UNC-5, possess 2 predicted Ig-like
 CC domains and 2 predicted thrombospondin type-1 repeats, a predicted
 CC membrane spanning region, and a large intracellular domain. They
 CC are predicted to be involved in cell migration and axon guidance,
 CC and are characterised as receptor proteins for netrins. Human
 CC UNC5H-1 (see AAW78899) and UNC5H-2 (see AAW78901) proteins are also
 CC claimed. Vertebrate UNC-5 proteins may be produced recombinantly
 CC from transfected host cells. The invention also provides unc-5
 CC hybridisation probes and primers, vertebrate UNC-5-specific binding
 CC agents such as specific antibodies, and methods of making and using
 CC the subject compositions in diagnosis (e.g. genetic hybridisation
 CC screens for vertebrate unc-5 transcripts), therapy (e.g. gene
 CC therapy to modulate vertebrate unc-5 gene expression) and in the
 CC biopharmaceutical industry (e.g. as immunogens, reagents for
 CC modulating cell guidance, reagents for screening chemical libraries
 CC for lead pharmacological agents, etc.).

XX CC (Updated on 25-MAR-2003 to correct PI field.)

XX SQ Sequence 898 AA;

Query Match 2.0%; Score 8; DB 19; Length 898;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 LGQLAAAV 348

| | | | |

Db 873 LGQLAAAV 880

| | | | |

RESULT 64

AAU97900
ID AAU97900 standard; Protein; 898 AA.
XX
AC AAU97900;
XX
DT 27-AUG-2002 (first entry)
XX
DE Rat netrin binding membrane receptor UNC5H-1 protein.
XX
KW Netrin binding membrane receptor; receptor; UNC5H-1;
KW Rat; neurotropic; neuroprotective; cytosolic; antiparkinsonian;
KW cerebroprotective; cancer; central nervous system; CNS; stroke;
KW Parkinson's disease; multiple sclerosis; Alzheimer's disease.
XX
OS Rattus sp.
XX
FH Key Location/Qualifiers
FT Domain 152..223
FT /note= "Immunoglobulin domain"
FT Domain 247..294
FT /note= "Thrombospondine type 1 domain"
FT Domain 302..348
FT /note= "Thrombospondine type 1 domain"
FT Region 361..382
FT /note= "Transmembrane region"
FT Domain 495..598
FT /note= "ZUS domain"
FT Domain 817..897
FT /note= "Death domain"
XX
PN WO200233080-A2.
XX
PD 25-APR-2002.
XX
PF 15-OCT-2001; 2001WO-EP11891.
XX
PR 16-OCT-2000; 2000US-240061P.
XX
PA (FARB) BAYER AG.
XX
PI Koehler RH;
XX
DR WPI; 2002-463314/49.
XX
PT Novel human netrin binding membrane receptor polypeptide and
PT polynucleotides for identifying modulating agents useful in treating
PT diseases e.g. Parkinson's disease, multiple sclerosis, stroke,
PT Alzheimer's disease
XX
PS Disclosure; Fig 3; 94pp; English.
XX
CC This invention relates to the DNA and protein sequences of a novel
CC purified human netrin binding membrane receptor, UNC5H-1. The DNA
CC sequence of the invention is useful as a probe for detecting a
CC nucleic acid encoding the UNC5H-1 protein in a biological sample.
CC The sequences of the invention are useful to screen for agents which
CC decrease the activity of the UNC5H-1 protein. The sequences are also
CC useful for screening agents which regulate (modulate) the activity of
CC the protein of the invention. A pharmaceutical composition containing
CC the protein of the invention or a reagent that modulates the activity
CC of the UNC5H-1 protein may be useful for treating a UNC5H-1 dysfunction
CC related disease such as cancer or a central nervous system (CNS)
CC disorders (e.g. Parkinson's disease, multiple sclerosis, stroke and
CC Alzheimer's disease). Fusion proteins comprising the UNC5H-1 protein are
CC useful for generating antibodies and for in various assay systems, and
CC the protein can be used as a bait protein in a two-hybrid assay or
CC three-hybrid assay. The method of the invention is useful for detecting
CC a coding sequence for the UNC5H-1 protein. The present sequence
CC represents the Rat netrin binding membrane receptor UNC5H-1 protein of
CC the invention.

SQ Sequence 898 AA;

Query Match 2.0%; Score 8; DB 23; Length 898;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 LQQLAAAV 348
DB 873 LQQLAAAV 880
|||||||

RESULT 65

AAU10543
ID AAU10543 standard; Protein; 898 AA.

XX
AC AAU10543;

XX
DT 14-FEB-2002 (first entry)

XX
DE Rat netrin receptor UNC5H1 (YSG7) polypeptide.

XX
KW YSG; YSG7; schizophrenia; chronic animal model; LCGU; netrin receptor;
KW local cerebral glucose utilisation; phosphodiesterase 1-alpha; UNC5H1;
KW calcium-independent alpha-latrotoxin receptor; CIRL; trkE; synapsin 1A;
KW epithelial discoidin domain receptor 1; synapsin 1B; neuroleptic;
KW tumour necrosis factor alpha; TNF-alpha; rat.

XX
OS Rattus sp.

XX
PN WO200175440-A2.

XX
PD 11-OCT-2001.

XX
PF 02-APR-2001; 2001WO-GB01486.

XX
PR 31-MAR-2000; 2000GB-0007880.

XX
PR 26-MAY-2000; 2000GB-0012768.

XX
PA (WELF-) WELFIDE CORP.

XX
PI Cochran S, Paterson G, Ohashi Y, Morris B, Pratt J;

XX
DR WPI; 2002-010813/01.

XX
DR N-PSDB; AAS16843.

XX
PT Novel chronic animal model of schizophrenia, useful for identifying
PT anti-psychotic drugs and genes that are associated with schizophrenia

XX
PS Disclosure; Fig 8b; 79pp; English.

XX
CC The invention relates to YSG polynucleotide fragments for use in
CC diagnosing and/or developing treatments for schizophrenia using chronic
CC animal models. The polynucleotides and their encoded polypeptides are
CC used for identification of compounds which modulate the expression of YSG
CC molecules, leading to the manufacture of schizophrenia medicaments. The
CC sequences can also be used for testing candidate compounds for any effect
CC on the polypeptides. Anti-schizophrenic effects of a compound can be
CC determined by measuring local cerebral glucose utilisation (LCGU) or
CC comparing its expression level with that of a control group. The
CC sequences are useful in the identification of genes associated with
CC schizophrenic states and in the development of an antibody. The sequences
CC of the invention include phosphodiesterase 1-alpha, calcium-independent
CC alpha-latrotoxin receptors (CIRL)-1-263, epithelial discoidin domain
CC receptor 1 (trkE), netrin receptor (UNC5H1), synapsins 1A and AB and
CC tumour necrosis factor (TNF) alpha. This sequence represents rat netrin
CC receptor UNC5H1 (YSG7) polypeptide.

XX
SQ Sequence 898 AA;

Query Match 2.0%; Score 8; DB 23; Length 898;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 LGQLAAAV 348
 |||||
 Db 873 LGQLAAAV 880

RESULT 66

AAU34320
 ID AAU34320 standard; Protein; 2478 AA.

XX
 AC AAU34320;

DT 14-FEB-2002 (first entry)

XX Staphylococcus aureus cellular proliferation protein #596.

XX Antisense; prokaryotic cellular proliferation protein;

KW antibiotic; antibacterial; drug design.

XX Staphylococcus aureus.

XX WO200170955-A2.

PN 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207277P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

DR N-PSDB; AAS52179.

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids -

XX Example 3; Seq ID No 5816; 51lpp; English.

XX The invention relates to antisense inhibitors of genes essential to

CC prokaryotic cellular proliferation, their use in identifying the

CC genes, their use in the discovery of novel antibiotics, the essential

CC genes themselves and the encoded proteins. The prokaryotes used are

CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella

CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The

CC invention is also useful for the identification of potential new targets

CC for antibiotic development. The antisense nucleic acids can also be used

CC to identify proteins used in proliferation, to express these proteins,

CC and to obtain antibodies capable of binding to the expressed proteins.

CC The proteins can be used to screen compounds in rational drug discovery

CC programmes. The antisense nucleic acid sequence is also useful to screen

CC for homologous nucleic acids which are required for cell proliferation in

CC a wide variety of organisms. The present sequence represents an

CC essential prokaryotic cellular proliferation protein.

CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic

CC format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2478 AA;

SQ

Query Match 2.0%; Score 8; DB 22; Length 2478;

Best Local Similarity 100.0%; Pred. No. 6.6e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 AQPAAPVA 104

|||||

Db 161 AQPAAPVA 168

RESULT 67

AAU37374

ID AAU37374 standard; Protein; 2478 AA.

XX AAU37374;

DT 14-FEB-2002 (first entry)

XX Staphylococcus aureus cellular proliferation protein #1544.

XX Antisense; prokaryotic cellular proliferation protein;

KW antibiotic; antibacterial; drug design.

XX Staphylococcus aureus.

XX WO200170955-A2.

PN 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207277P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

DR N-PSDB; AAS55233.

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids -

XX Example 3; Seq ID No 12967; 51lpp; English.

CC The invention relates to antisense inhibitors of genes essential to

CC prokaryotic cellular proliferation, their use in identifying the

CC genes, their use in the discovery of novel antibiotics, the essential

CC genes themselves and the encoded proteins. The prokaryotes used are

CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella

CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The

CC invention is also useful for the identification of potential new targets

CC for antibiotic development. The antisense nucleic acids can also be used

CC to identify proteins used in proliferation, to express these proteins,

CC and to obtain antibodies capable of binding to the expressed proteins.

CC The proteins can be used to screen compounds in rational drug discovery

CC programmes. The antisense nucleic acid sequence is also useful to screen

CC for homologous nucleic acids which are required for cell proliferation in

CC a wide variety of organisms. The present sequence represents an

CC essential prokaryotic cellular proliferation protein.

CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic

CC format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2478 AA;

SQ

Query Match 2.0%; Score 8; DB 22; Length 2478;

Best Local Similarity 100.0%; Pred. No. 6.6e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 AQAAPVA 104
|||||||
Db 161 AQAAPVA 168

RESULT 68

ABJ19002
ID ABJ19002 standard; Protein; 2478 AA.

XX
AC ABJ19002;

XX
DT 06-MAR-2003 (first entry)

XX Pathogen specific antigen related staphylococcal protein SEQ ID No 220.

XX Antibacterial; virucide; fungicide; protozoicide; cytostatic; anti-HIV;
XX hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;
XX auto-immunity; vaccine; staphylococcal infection; antibody; cancer;
XX autoimmune disease; HIV; hepatitis.

OS Staphylococcus sp.

XX WO200259148-A2.

XX
PD 01-AUG-2002.

XX
PF 21-JAN-2002; 2002WO-EP00546.

XX
PR 26-JAN-2001; 2001AT-0000130.

XX (CIST-) CISTEM BIOTECHNOLOGIES GMBH.

XX
PI Meinke A, Nagy E, Von Ahnen U, Klade C, Henics T, Zauner W;
PI Minh DB, Vytvytska O, Etz H, Dryla A, Weichhart T, Hafner M;
PI Tempelmaier B;

XX WPI; 2003-075410/07.

XX Identifying, isolating and producing hyperimmune serum-reactive
PT antigens from a pathogen, for preparing vaccine or medicament for
PT treating or preventing e.g. staphylococcal infections, comprises
PT providing antibody preparation -

PS Example 7; Page 186; 252pp; English.

XX The invention relates to a novel method for identifying, isolating and
CC producing hyperimmune serum-reactive antigens from a pathogen, tumour,
CC allergen, a tissue or host prone to auto-immunity, where the antigens
CC are used in a vaccine, comprises providing antibody preparation from a
CC plasma pool of a type of animal, or individual sera with antibodies
CC against the specific pathogen, tumour, allergen, tissue or host prone to
CC auto-immunity. The hyperimmune serum-reactive antigens comprising any of
CC the 62 sequences of 53-2261 amino acids fully defined in the
CC specification, or their hyperimmune fragments are useful for the
CC manufacture of a pharmaceutical preparation, particularly a vaccine
CC against staphylococcal infections or colonisation against *S. aureus* or *S.*
CC epidermidis. The preparation of antibodies is useful for the manufacture
CC of a medicament for treating or preventing staphylococcal infections or
CC colonisation against *S. aureus* or *S. epidermidis*. The antibody
CC preparations may also be used for diagnostic and imaging purposes. Other
CC conditions that can be treated include cancer, autoimmune diseases or
CC infections caused by viral (e.g. HIV, hepatitis A, B or C), fungal or
CC protozoan pathogens. This sequence represents a staphylococcal protein
CC relating to the method for identifying and producing pathogen specific
CC antigens of the invention.

SQ Sequence 2478 AA;

Query Match 2.0%; Score 8; DB 24; Length 2478;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 AQAAPVA 104
|||||||
Db 161 AQAAPVA 168

RESULT 69

ABG96596
ID ABG96596 standard; Peptide; 9 AA.

XX
AC ABG96596;

XX
DT 16-DEC-2002 (first entry)

XX Human leukocyte antigen (HLA) B15 ligand #19.

XX Soluble human leukocyte antigen; HLA; SHLA; cell pharm;
KW multimeric HLA complex; cytotoxic T lymphocytes; CTL; immune response;
KW viral infection; cancer; autoimmune disease; vaccine development; MHC;
KW major histocompatibility complex; diagnostic development;
KW HLA class I polymorphism; HLA-B15 allotype; ligand.

OS Homo sapiens.

XX WO200262846-A2.

XX
PD 15-AUG-2002.

XX
PF 18-DEC-2001; 2001WO-US49744.

XX
PR 18-DEC-2000; 2000US-256409P.

XX
PR 18-DEC-2000; 2000US-256410P.

XX
PR 10-OCT-2001; 2001US-0974366.

XX (HILD/) HILDEBRAND W H.

XX (PRIL/) PRILLIMAN K R.

XX Hildebrand WH, Prilliman KR;

XX WPI; 2002-698563/75.

XX Producing soluble human leukocyte antigen (HLA) in cell pharm useful
PT for studies of peptide loading for characterizing human immune
PT responses involves using HLA allelic cDNA or genomic DNA as starting
PT material -

XX Disclosure; Page 159; 300pp; English.

XX The invention describes a method of producing soluble human leukocyte
CC antigen (HLA) molecules (sHLA) in cell pharm involving amplifying HLA
CC allelic DNA by PCR using a locus specific primer to produce truncated a
CC PCR product (PI), inserting PI into mammalian expression vector;
CC electroporating the plasmid into a host cell; inoculating the cell pharm
CC with the host cell such that cell pharm produces sHLA. A multimeric HLA
CC complex (I) is useful for testing functionality of peptide ligands bound
CC by at least two soluble HLA molecules. (I) can be tested for its ability
CC to serve as ligands for cytotoxic T lymphocytes (CTLs) and induce immune
CC responses in humans. (I) is useful for studying T cell responses to
CC pathological conditions such as viral infections and cancer, and for
CC modulating the human immune system to induce tolerance in autoimmune
CC diseases. The individual secreted major histocompatibility complex (MHC)
CC molecules produced are useful for studies of peptide loading (i.e., in
CC vaccine development) and to the development of diagnostics with the
CC secreted MHC molecules, naturally loaded peptides can be eluted from the
CC MHC molecule and characterised. The secreted MHC molecules allow the
CC assessment of structural and functional impact of HLA class I
CC polymorphism. The molecules are also useful to generate ligands and hence
CC ligand maps from the peptide pools extracted from series of distinct yet
CC related class I HLA-B15 allotypes; compare the different ligand maps to
CC identify potentially shared elements; and characterise the elements
CC identified to positively or negatively validate the occurrence of
CC overlapping ligands. The truncated version of (MHC) can be produced in
CC mammalian or insect/bacterial cells such that milligram or greater
CC quantities of an individual class I or class II molecule can be obtained.

CC This sequence represents a HLA (human leukocyte antigen) peptide
 CC ligand.
 XX
 SQ Sequence 9 AA;
 Query Match 1.7%; Score 7; DB 23; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 284 QFQALLQ 290
 Db 2 QFQALLQ 8
 RESULT 70
 ABG96636
 ID ABG96636 standard; Peptide; 9 AA.
 XX
 AC ABG96636;
 XX
 DT 16-DEC-2002 (first entry)
 XX
 DE Human leukocyte antigen (HLA) B15 ligand #59.
 XX
 KW Soluble human leukocyte antigen; HLA; sHLA; cell pharm;
 KW multimeric HLA complex; cytotoxic T lymphocytes; CTL; immune response;
 KW viral infection; cancer; autoimmune disease; vaccine development; MHC;
 KW major histocompatibility complex; diagnostic development;
 KW HLA class I polymorphism; HLA-B15 allotype; ligand.
 XX
 OS Homo sapiens.
 XX
 PN WO200262846-A2.
 XX
 PD 15-AUG-2002.
 XX
 PF 18-DEC-2001; 2001WO-US49744.
 XX
 PR 18-DEC-2000; 2000US-256409P.
 PR 18-DEC-2000; 2000US-256410P.
 PR 10-OCT-2001; 2001US-0974366.
 XX
 PA (HILD/) HILDEBRAND W H.
 PA (PRIL/) PRILLIMAN K R.
 XX
 PI Hildebrand WH, Prilliman KR;
 XX
 DR WPI; 2002-698563/75.
 XX
 PT Producing soluble human leukocyte antigen (HLA) in cell pharm useful
 PT for studies of peptide loading for characterizing human immune
 PT responses involves using HLA allelic cDNA or genomic DNA as starting
 PT material -
 XX
 PS Disclosure; Page 161; 300pp; English.
 XX
 CC The invention describes a method of producing soluble human leukocyte
 CC antigen (HLA) molecules (sHLA) in cell pharm involving amplifying HLA
 CC allelic DNA by PCR using a locus specific primer to produce truncated a
 CC PCR product (PI), inserting PI into mammalian expression vector;
 CC electroporating the plasmid into a host cell; inoculating the cell pharm
 CC with the host cell such that cell pharm produces sHLA. A multimeric HLA
 CC complex (I) is useful for testing functionality of peptide ligands bound
 CC by at least two soluble HLA molecules. (I) can be tested for its ability
 CC to serve as ligands for cytotoxic T lymphocytes (CTLs) and induce immune
 CC responses in humans. (I) is useful for studying T cell responses to
 CC pathological conditions such as viral infections and cancer, and for
 CC modulating the human immune system to induce tolerance in autoimmune
 CC diseases. The individual secreted major histocompatibility complex (MHC)
 CC molecules produced are useful for studies of peptide loading (i.e., in
 CC vaccine development) and to the development of diagnostics. With the
 CC secreted MHC molecules, naturally loaded peptides can be eluted from the
 CC MHC molecule and characterised. The secreted MHC molecules allow the

CC assessment of structural and functional impact of HLA class I
 CC polymorphism. The molecules are also useful to generate ligands and hence
 CC ligand maps from the peptide pools extracted from series of distinct yet
 CC related class I HLA-B15 allotypes; compare the different ligand maps to
 CC identify potentially shared elements; and characterise the elements
 CC identified to positively or negatively validate the occurrence of
 CC overlapping ligands. The truncated version of (MHC) can be produced in
 CC mammalian or insect/bacterial cells such that milligram or greater
 CC quantities of an individual class I or class II molecule can be obtained.
 CC This sequence represents a HLA (human leukocyte antigen) peptide
 CC ligand.
 XX
 SQ Sequence 9 AA;
 Query Match 1.7%; Score 7; DB 23; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 284 QFQALLQ 290
 Db 2 QFQALLQ 8
 RESULT 71
 ABG97255
 ID ABG97255 standard; Peptide; 9 AA.
 XX
 AC ABG97255;
 XX
 DT 16-DEC-2002 (first entry)
 XX
 DE Human leukocyte antigen (HLA) B15 ligand #584.
 XX
 KW Soluble human leukocyte antigen; HLA; sHLA; cell pharm;
 KW multimeric HLA complex; cytotoxic T lymphocytes; CTL; immune response;
 KW viral infection; cancer; autoimmune disease; vaccine development; MHC;
 KW major histocompatibility complex; diagnostic development;
 KW HLA class I polymorphism; HLA-B15 allotype; ligand.
 XX
 OS Homo sapiens.
 XX
 PN WO200262846-A2.
 XX
 PD 15-AUG-2002.
 XX
 PF 18-DEC-2001; 2001WO-US49744.
 XX
 PR 18-DEC-2000; 2000US-256409P.
 PR 18-DEC-2000; 2000US-256410P.
 PR 10-OCT-2001; 2001US-0974366.
 XX
 PA (HILD/) HILDEBRAND W H.
 PA (PRIL/) PRILLIMAN K R.
 XX
 PI Hildebrand WH, Prilliman KR;
 XX
 DR WPI; 2002-698563/75.
 XX
 PT Producing soluble human leukocyte antigen (HLA) in cell pharm useful
 PT for studies of peptide loading for characterizing human immune
 PT responses involves using HLA allelic cDNA or genomic DNA as starting
 PT material -
 XX
 PS Disclosure; Page 188; 300pp; English.
 XX
 CC The invention describes a method of producing soluble human leukocyte
 CC antigen (HLA) molecules (sHLA) in cell pharm involving amplifying HLA
 CC allelic DNA by PCR using a locus specific primer to produce truncated a
 CC PCR product (PI), inserting PI into mammalian expression vector;
 CC electroporating the plasmid into a host cell; inoculating the cell pharm
 CC with the host cell such that cell pharm produces sHLA. A multimeric HLA
 CC complex (I) is useful for testing functionality of peptide ligands bound
 CC by at least two soluble HLA molecules. (I) can be tested for its ability

CC to serve as ligands for cytotoxic T lymphocytes (CTLs) and induce immune
 CC responses in humans. (I) is useful for studying T cell responses to
 CC pathological conditions such as viral infections and cancer, and for
 CC modulating the human immune system to induce tolerance in autoimmune
 CC diseases. The individual secreted major histocompatibility complex (MHC)
 CC molecules produced are useful for studies of peptide loading (i.e., in
 CC vaccine development) and to the development of diagnostics. With the
 CC secreted MHC molecules, naturally loaded peptides can be eluted from the
 CC MHC molecule and characterised. The secreted MHC molecules allow the
 CC assessment of structural and functional impact of HLA class I
 CC polymorphism. The molecules are also useful to generate ligands and hence
 CC ligand maps from the peptide pools extracted from series of distinct yet
 CC related class I HLA-B*57 allotypes; compare the different ligand maps to
 CC identify potentially shared elements; and characterise the elements
 CC identified to positively or negatively validate the occurrence of
 CC overlapping ligands. The truncated version of (MHC) can be produced in
 CC mammalian or insect/bacterial cells such that milligram or greater
 CC quantities of an individual class I or class II molecule can be obtained.
 CC This sequence represents a HLA (human leukocyte antigen) peptide
 CC ligand.

XX Sequence 9 AA;
 SQ

Query Match 1.7%; Score 7; DB 23; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 284 QFQALLQ 290
 Db 2 QFQALLQ 8

RESULT 72

AAG94101

ID AAG94101 standard; Peptide; 10 AA.

AC AAG94101;

DT 18-SEP-2001 (first entry)

XX Human complementary peptide, SEQ ID NO: 295.

DE Human complementary peptide; ligand; drug discovery; drug design.

KW Homo sapiens.

XX WO200142277-A2.

PN 14-JUN-2001.

XX 13-DEC-2000; 2000WO-GB04776.

XX 13-DEC-1999; 99GB-0029464.

XX (PROT-) PROTEOM LTD.

XX Roberts GW, Heal JR;

XX WPI; 2001-408419/43.

XX A set of peptide ligands consisting of specific complementary peptides
 PT to proteins encoded by genes of the human genome, useful in an assay
 PT for screening and identifying of one or more novel peptides which are
 PT drug candidates or pro-drugs -

XX Example 4; Page 82; 646pp; English.

XX The invention relates to a set of complementary peptide ligands
 CC generated from the human genome. The complementary peptides
 CC interact with their relevant target proteins encoded in the human
 CC genome. They can be used as reagents in drug discovery and as lead
 CC ligands to facilitate drug design and development. The present
 CC sequence is a complementary peptide provided in the specification.

XX Sequence 10 AA;
 SQ

Query Match 1.7%; Score 7; DB 22; Length 10;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 PAASVAR 111
 Db 1 PAASVAR 7

RESULT 73

ABG02484

ID ABG02484 standard; Protein; 52 AA.

AC ABG02484;

DT 13-FEB-2002 (first entry)

XX Novel human diagnostic protein #2475.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS66671.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

PS Claim 20; SEQ ID No 32843; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (III). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 52 AA;
 SQ

Query Match 1.7%; Score 7; DB 22; Length 52;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 PQAPAA 136
|||||
Db 3 PQAPAA 9

RESULT 74

ABG57158
ID ABG57158 standard; Peptide; 60 AA.

XX AC ABG57158;

XX DT 25-FEB-2003 (first entry)

XX DE Human liver peptide, SEQ ID No 35806.

XX KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
hypercholesterolaemia; coronary heart disease.

XX OS Homo sapiens.

XX PN WO200157273-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00664.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488898/53.

XX PT Human genome-derived single exon nucleic acid probes useful for
analyzing gene expression in human adult liver -

XX PS Claim 27; SEQ ID No 35806; 658pp; English.

XX CC The invention relates to a single exon nucleic acid probe (SEN) (I) for
measuring human gene expression in a sample derived from human adult
liver, comprising one of 13109 defined nucleotide sequences given in the
specification (or complements/ fragments). The probe hybridises at high
stringency to a nucleic acid molecule expressed in the human adult
liver. (I) may be used for predicting, measuring and displaying gene
expression in samples derived from human adult liver. The genes
identified may be involved in genetic liver diseases such as cirrhosis,
hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
is associated with coronary heart disease. ABG47348-ABG59930 represent
human liver single exon encoded peptides of the invention.

XX CC Note: The sequence information for this patent does not appear in the
printed specification but was obtained in electronic format directly
from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 60 AA;

Query Match 1.7%; Score 7; DB 22; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 PVAPAA 108
|||||

Db 42 PVAPAA 48

RESULT 75

ABB41719
ID ABB41719 standard; Peptide; 60 AA.

XX AC ABB41719;

XX DT 04-FEB-2002 (first entry)

XX DE Peptide #9225 encoded by human foetal liver single exon probe.

XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.

XX OS Homo sapiens.

XX PN WO200157277-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00669.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-483447/52.

XX PT Human genome-derived single exon nucleic acid probes useful for
analyzing gene expression in human fetal liver -

XX PS Claim 27; SEQ ID No 34354; 639pp + sequence listing; English.

XX CC The invention relates to a single exon nucleic acid probe for
measuring human gene expression in a sample derived from human foetal
liver. The single exon nucleic acid probes may be used for predicting,
measuring and displaying gene expression in samples derived from human
fetal liver. The present sequence is a peptide encoded by a single exon
nucleic acid probe of the invention.

XX CC Note: The sequence data for this patent did not form part of the
printed specification, but was obtained in electronic format directly
from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 60 AA;

Query Match 1.7%; Score 7; DB 22; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 PVAPAA 108
|||||

Db 42 PVAPAA 48

Search completed: December 17, 2003, 06:33:18
Job time : 42.2471 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 06:35:27 ; Search time 34.0556 Seconds
(without alignments)
2220.991 Million cell updates/sec

Title: US-09-805-550-2

Perfect score: 405

Sequence: 1 MKLNVKTLKGTNFEIASPD.....EELTANYLLDHGHEFDQQQ 405

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 696363 seqs, 186758610 residues

Word size :

Total number of hits satisfying chosen parameters: 6963363

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/prodata/1/pubpaa/us07_PUBCOMB.pcp.*
- 2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pcp.*
- 3: /cgn2_6/prodata/1/pubpaa/us06_NEW_PUB.pcp.*
- 4: /cgn2_6/prodata/1/pubpaa/us06_PUBCOMB.pcp.*
- 5: /cgn2_6/prodata/1/pubpaa/us07_NEW_PUB.pcp.*
- 6: /cgn2_6/prodata/1/pubpaa/PCTus_PUBCOMB.pcp.*
- 7: /cgn2_6/prodata/1/pubpaa/us08_NEW_PUB.pcp.*
- 8: /cgn2_6/prodata/1/pubpaa/us08_PUBCOMB.pcp.*
- 9: /cgn2_6/prodata/1/pubpaa/us09A_PUBCOMB.pcp.*
- 10: /cgn2_6/prodata/1/pubpaa/us09B_PUBCOMB.pcp.*
- 11: /cgn2_6/prodata/1/pubpaa/us09C_PUBCOMB.pcp.*
- 12: /cgn2_6/prodata/1/pubpaa/us09_NEW_PUB.pcp.*
- 13: /cgn2_6/prodata/1/pubpaa/us10A_PUBCOMB.pcp.*
- 14: /cgn2_6/prodata/1/pubpaa/us10B_PUBCOMB.pcp.*
- 15: /cgn2_6/prodata/1/pubpaa/us10C_PUBCOMB.pcp.*
- 16: /cgn2_6/prodata/1/pubpaa/us10_NEW_PUB.pcp.*
- 17: /cgn2_6/prodata/1/pubpaa/us60_NEW_PUB.pcp.*
- 18: /cgn2_6/prodata/1/pubpaa/us60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Query			Description
			Match	Length	ID	
1	405	100.0	405	9	US-09-805-550-2	Sequence 2, Appli
2	13	3	368	9	US-09-805-550-4	Sequence 4, Appli
3	9	2.2	3069	10	US-09-712-363-246	Sequence 246, App
4	8	2.0	186	9	US-09-729-674-30	Sequence 30, Appl
5	8	2.0	227	15	US-10-156-761-11941	Sequence 11941, A
6	8	2.0	254	9	US-09-815-242-13859	Sequence 13859, A
7	8	2.0	540	8	US-08-781-986A-5245	Sequence 5225, Ap
8	8	2.0	540	8	US-08-781-986A-5245	Sequence 5242, Ap
9	8	2.0	596	11	US-09-946-374-243	Sequence 243, App
10	8	2.0	596	12	US-10-015-387A-243	Sequence 243, App
11	8	2.0	596	12	US-10-063-735-100	Sequence 100, App
12	8	2.0	596	12	US-10-006-130A-243	Sequence 243, App
13	8	2.0	596	12	US-10-199-672-310	Sequence 310, App
14	8	2.0	596	12	US-10-006-172A-243	Sequence 243, App
15	8	2.0	596	12	US-10-187-749-310	Sequence 310, App

```
89      8      2.0      596      12      US-10-063-583-100      Sequence 100, App
90      8      2.0      596      12      US-10-063-584-100      Sequence 100, App
91      8      2.0      596      12      US-10-063-587-100      Sequence 100, App
92      8      2.0      596      12      US-10-063-589-100      Sequence 100, App
93      8      2.0      596      12      US-10-063-591-100      Sequence 100, App
94      8      2.0      596      12      US-10-063-592-100      Sequence 100, App
95      8      2.0      596      12      US-10-063-593-100      Sequence 100, App
96      8      2.0      596      12      US-10-063-596-100      Sequence 100, App
97      8      2.0      596      12      US-10-063-597-100      Sequence 100, App
98      8      2.0      596      12      US-10-063-600-100      Sequence 100, App
99      8      2.0      596      12      US-10-063-602-100      Sequence 100, App
100     8      2.0      596      12      US-10-063-604-100      Sequence 100, App
```

ALIGNMENTS

```
RESULT 1
US-09-805-550-2
; Sequence 2, Application US/09805550
; Patent No. US2002026045A1
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; APPLICANT: Tagliani, Laura
; TITLE OF INVENTION: Rad23 Genes and Uses Thereof
; FILE REFERENCE: 0964D
; CURRENT APPLICATION NUMBER: US/09/805,550
; CURRENT FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 09/413,574
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: 60/109,728
; PRIOR FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Zea mays
US-09-805-550-2
```

```
Query Match      100.0%; Score 405; DB 9; Length 405;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1      M K L N V K T L K G T N F E I E A S P D A S V A D V K E I I E T T O G S T Y R A D Q O M L I Y O G K I L K D E T T L E      60
Db      1      M K L N V K T L K G T N F E I E A S P D A S V A D V K E I I E T T O G S T Y R A D Q O M L I Y O G K I L K D E T T L E      60

Qy      61      S N G V A E N S F L V I M L S K A K A S S G A S T A T T A K A P A T L A Q P A A P V A P A A S V A R T P T Q A P V A T      120
Db      61      S N G V A E N S F L V I M L S K A K A S S G A S T A T T A K A P A T L A Q P A A P V A P A A S V A R T P T Q A P V A T      120

Qy      121     A E T A P S V Q P Q A P A A T V A A T D D A D V Y S Q A A S N L V F G N N L E Q T I Q O I L D M G G G T W E R D T V      180
Db      121     A E T A P S V Q P Q A P A A T V A A T D D A D V Y S Q A A S N L V F G N N L E Q T I Q O I L D M G G G T W E R D T V      180

Qy      181     V R A L P A A Y N N P E R A I D Y L S G I P E N V E A Q P V A P A A C Q Q T N Q Q A S P A Q P A V A L P V Q P S      240
Db      181     V R A L P A A Y N N P E R A I D Y L S G I P E N V E A Q P V A P A A C Q Q T N Q Q A S P A Q P A V A L P V Q P S      240

Qy      241     P A S A G P N A P N L P F Q G V P S G S N P G V P G A G S G A L D A L R O L P Q F A L L Q L V Q A N P Q I L Q      300
Db      241     P A S A G P N A P N L P F Q G V P S G S N P G V P G A G S G A L D A L R O L P Q F A L L Q L V Q A N P Q I L Q      300

Qy      301     P M L Q E L G K Q N P Q I L R I Q E N Q A E F L R L V N E S P E G P G G N I L G Q L A A A V P Q T L T V T P E E R E      360
Db      301     P M L Q E L G K Q N P Q I L R I Q E N Q A E F L R L V N E S P E G P G G N I L G Q L A A A V P Q T L T V T P E E R E      360

Qy      361     A I Q R L E G G F N R E L V L E V F F A C N K D E E L T A N Y L L D H G H E F D D Q Q      405
Db      361     A I Q R L E G G F N R E L V L E V F F A C N K D E E L T A N Y L L D H G H E F D D Q Q      405
```

```
RESULT 2
US-09-805-550-4
; Sequence 4, Application US/09805550
; Patent No. US2002026045A1
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; APPLICANT: Tagliani, Laura
; TITLE OF INVENTION: Rad23 Genes and Uses Thereof
; FILE REFERENCE: 0964D
; CURRENT APPLICATION NUMBER: US/09/805,550
; CURRENT FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 09/413,574
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: 60/109,728
; PRIOR FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Zea mays
US-09-805-550-4
```

```
Query Match      3.2%; Score 13; DB 9; Length 368;
Best Local Similarity 100.0%; Pred. No. 0.0023;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      182     R A L R A A Y N N P E R A      194
Db      167     R A L R A A Y N N P E R A      179
```

```
RESULT 3
US-09-712-363-246
; Sequence 246, Application US/09712363
; Patent No. US20020164588A1
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, David
; APPLICANT: Rotstein, Edward M.
; APPLICANT: Marcotte, Sergio H.
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
; TITLE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
; FILE REFERENCE: 07419-032001
; CURRENT APPLICATION NUMBER: US/09/712,363
; CURRENT FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: PCT/US00/02246
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,531
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/117,844
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 246
; LENGTH: 3069
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-246
```

```
Query Match      2.2%; Score 9; DB 10; Length 3069;
Best Local Similarity 100.0%; Pred. No. 64;
```

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 AAPVAPAA 108
| | | | | | | | | |
Db 1754 AAPVAPAA 1762

RESULT 4

US-09-729-674-30
; Sequence 30, Application US/09729674
; Patent No. US2001003935A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Pechtel, Kim
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6053-64X
; CURRENT APPLICATION NUMBER: US/09/729,674
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/539,330
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 283
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 30
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-729-674-30

Query Match 2.0%; Score 8; DB 9; Length 186;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
| | | | | | | | | |
Db 39 SSGASTAT 46

RESULT 5

US-10-156-761-11941
; Sequence 11941, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11941
; LENGTH: 227
; TYPE: PRT

; ORGANISM: Streptomyces avermitilis
US-10-156-761-11941

Query Match 2.0%; Score 8; DB 15; Length 227;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 AAPVAPAA 107
| | | | | | | | | |
Db 30 AAPVAPAA 37

RESULT 6

US-09-815-242-13899
; Sequence 13899, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13899
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Salmonella typhi
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(254)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-13899

Query Match 2.0%; Score 8; DB 9; Length 254;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 REAIQRL 366
| | | | | | | | | |
Db 52 REAIQRL 59

RESULT 7

US-08-781-986A-5225
; Sequence 5225, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 5225:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 540 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-781-986A-5225

Query Match 2.0%; Score 8; DB 8; Length 540;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 AQPAPVA 104
Db 175 AQPAPVA 182

RESULT 8
US-08-781-986A-5242
; Sequence 5242, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob

; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 5242:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 540 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-781-986A-5242

Query Match 2.0%; Score 8; DB 8; Length 540;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 AQPAPVA 104
Db 175 AQPAPVA 182

RESULT 9
US-09-946-374-243
; Sequence 243, Application US/09946374
; Publication No. US20030073129A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830F1C1
; CURRENT APPLICATION NUMBER: US/09/946,374
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/098716
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098723
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098749
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098750
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098803
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098821
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098843
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/099536
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099596
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099598

1	PRIOR APPLICATION NUMBER: 60/101479	
2	PRIOR FILING DATE: 1998-09-23	
3	PRIOR APPLICATION NUMBER: 60/101738	
4	PRIOR FILING DATE: 1998-09-24	
5	PRIOR APPLICATION NUMBER: 60/101741	
6	PRIOR FILING DATE: 1998-09-24	
7	PRIOR APPLICATION NUMBER: 60/101743	
8	PRIOR FILING DATE: 1998-09-24	
9	PRIOR APPLICATION NUMBER: 60/101915	
10	PRIOR FILING DATE: 1998-09-24	
11	PRIOR APPLICATION NUMBER: 60/101916	
12	PRIOR FILING DATE: 1998-09-24	
13	PRIOR APPLICATION NUMBER: 60/102207	
14	PRIOR FILING DATE: 1998-09-29	
15	PRIOR APPLICATION NUMBER: 60/102240	
16	PRIOR FILING DATE: 1998-09-29	
17	PRIOR APPLICATION NUMBER: 60/102307	
18	PRIOR FILING DATE: 1998-09-29	
19	PRIOR APPLICATION NUMBER: 60/102330	
20	PRIOR FILING DATE: 1998-09-29	
21	PRIOR APPLICATION NUMBER: 60/102331	
22	PRIOR FILING DATE: 1998-09-29	
23	PRIOR APPLICATION NUMBER: 60/102484	
24	PRIOR FILING DATE: 1998-09-30	
25	PRIOR APPLICATION NUMBER: 60/102487	
26	PRIOR FILING DATE: 1998-09-30	
27	PRIOR APPLICATION NUMBER: 60/102570	
28	PRIOR FILING DATE: 1998-09-30	
29	PRIOR APPLICATION NUMBER: 60/102571	
30	PRIOR FILING DATE: 1998-09-30	
31	PRIOR APPLICATION NUMBER: 60/102684	
32	PRIOR FILING DATE: 1998-10-01	
33	PRIOR APPLICATION NUMBER: 60/102687	
34	PRIOR FILING DATE: 1998-10-01	
35	PRIOR APPLICATION NUMBER: 60/102965	
36	PRIOR FILING DATE: 1998-10-02	
37	PRIOR APPLICATION NUMBER: 60/103258	
38	PRIOR FILING DATE: 1998-10-06	
39	PRIOR APPLICATION NUMBER: 60/103314	
40	PRIOR FILING DATE: 1998-10-07	
41	PRIOR APPLICATION NUMBER: 60/103315	
42	PRIOR FILING DATE: 1998-10-07	
43	PRIOR APPLICATION NUMBER: 60/103328	
44	PRIOR FILING DATE: 1998-10-07	
45	PRIOR APPLICATION NUMBER: 60/103395	
46	PRIOR FILING DATE: 1998-10-07	
47	PRIOR APPLICATION NUMBER: 60/103396	
48	PRIOR FILING DATE: 1998-10-07	
49	PRIOR APPLICATION NUMBER: 60/103401	
50	PRIOR FILING DATE: 1998-10-07	
51	PRIOR APPLICATION NUMBER: 60/103449	
52	PRIOR FILING DATE: 1998-10-06	
53	PRIOR APPLICATION NUMBER: 60/103633	
54	PRIOR FILING DATE: 1998-10-08	
55	PRIOR APPLICATION NUMBER: 60/103678	
56	PRIOR FILING DATE: 1998-10-08	
57	PRIOR APPLICATION NUMBER: 60/103679	
58	PRIOR FILING DATE: 1998-10-08	
59	PRIOR APPLICATION NUMBER: 60/103711	
60	PRIOR FILING DATE: 1998-10-08	
61	PRIOR APPLICATION NUMBER: 60/104257	
62	PRIOR FILING DATE: 1998-10-14	
63	PRIOR APPLICATION NUMBER: 60/104987	
64	PRIOR FILING DATE: 1998-10-20	
65	PRIOR APPLICATION NUMBER: 60/105000	
66	PRIOR FILING DATE: 1998-10-20	
67	PRIOR APPLICATION NUMBER: 60/105002	
68	PRIOR FILING DATE: 1998-10-20	
69	PRIOR APPLICATION NUMBER: 60/105104	
70	PRIOR FILING DATE: 1998-10-21	
71	PRIOR APPLICATION NUMBER: 60/105169	
72	PRIOR FILING DATE: 1998-10-22	
73	PRIOR APPLICATION NUMBER: 60/105266	

; PRIOR FILING DATE: 1998-10-22
 ; PRIOR APPLICATION NUMBER: 60/105693
 ; PRIOR FILING DATE: 1998-10-26
 ; PRIOR APPLICATION NUMBER: 60/105694
 ; PRIOR FILING DATE: 1998-10-26
 ; PRIOR APPLICATION NUMBER: 60/105807

Query Match 2.0%; Score 8; DB 11; Length 596;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
 Db 39 SSGASTAT 46

RESULT 10

US-10-015-387A-243
 ; Sequence 243, Application US/10015387A
 ; Publication No. US20030135034A1

GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan I.
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Hillan, Kenneth J.
 ; APPLICANT: Pan, James
 ; APPLICANT: Paoni, Nicholas F.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: P2830P1C54
 ; CURRENT APPLICATION NUMBER: US/10/015,387A

; PRIOR FILING DATE: 2001-12-12
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 477

; SEQ ID NO 243
 ; LENGTH: 596
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-015-387A-243

Query Match 2.0%; Score 8; DB 12; Length 596;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
 Db 39 SSGASTAT 46

RESULT 11

US-10-063-735-100
 ; Sequence 100, Application US/10063735
 ; Publication No. US20030138882A1

GENERAL INFORMATION:

; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; TITLE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P3230R1C1
 ; CURRENT APPLICATION NUMBER: US/10/063,735
 ; CURRENT FILING DATE: 2002-05-08

; Prior Application removed - See Palm or File Wrapper
 ; NUMBER OF SEQ ID NOS: 170
 ; SEQ ID NO 100
 ; LENGTH: 596

; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-063-735-100

Query Match 2.0%; Score 8; DB 12; Length 596;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
 Db 39 SSGASTAT 46

RESULT 12

US-10-006-130A-243
 ; Sequence 243, Application US/10006130A
 ; Publication No. US20030148375A1

GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan I.
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Hillan, Kenneth J.
 ; APPLICANT: Pan, James
 ; APPLICANT: Paoni, Nicholas F.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; TITLE OF INVENTION: Acids Encoding the Same
 ; FILE REFERENCE: P2830P1C7
 ; CURRENT APPLICATION NUMBER: US/10/006,130A

; PRIOR FILING DATE: 2002-03-19
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 477

; SEQ ID NO 243
 ; LENGTH: 596
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-006-130A-243

Query Match 2.0%; Score 8; DB 12; Length 596;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
 Db 39 SSGASTAT 46

RESULT 13

US-10-199-672-310
 ; Sequence 310, Application US/10199672
 ; Publication No. US20030148442A1

GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James

```

; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC1
; CURRENT APPLICATION NUMBER: US/10/199,672
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US/10/052,586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-199-672-310

Query Match      2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      81 SSGASTAT 88
DB      39 SSGASTAT 46
|||||||

RESULT 14
US-10-006-172A-243
; Sequence 243, Application US/10006172A
; Publication No. US20030153000A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830PIC11
; CURRENT APPLICATION NUMBER: US/10/006,172A
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/098716
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098723
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/098749
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098750
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098803
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098821
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098843
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/099536
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099596
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099598
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099602
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099642
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099741
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099754
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099763
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099792
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099808
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099815
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099816
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100385
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100388
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100390
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100584
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100627
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100661
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100662
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100664
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100683
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100684
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100710
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100711
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100848
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/100849
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/100919
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100930
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/101014
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/101068
; PRIOR FILING DATE: 1998-09-18

```

; PRIOR APPLICATION NUMBER: 60/101071
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/101279
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: 60/101471
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101472
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101474
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101475
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101476
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101477
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101479
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101738
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101741
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101743
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101915
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101916
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/102207
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/102240
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/102307
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/102330
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/102331
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/102484
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 60/102487
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 60/102570
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 60/102571
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 60/102684
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 60/102687
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 60/102965
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 60/103258
; PRIOR FILING DATE: 1998-10-06
; PRIOR APPLICATION NUMBER: 60/103314
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103315
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103328
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103395
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103396
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103401
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103449
; PRIOR FILING DATE: 1998-10-06
; PRIOR APPLICATION NUMBER: 60/103633
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/103678
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/103679

; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/103711
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/104257
; PRIOR FILING DATE: 1998-10-14
; PRIOR APPLICATION NUMBER: 60/104987
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105000
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105002
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105104
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 60/105169
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: 60/105266
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: 60/105693
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: 60/105694
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: 60/105807
; PRIOR FILING DATE: 1998-10-27
; PRIOR APPLICATION NUMBER: 60/105881
; PRIOR FILING DATE: 1998-10-27
; PRIOR APPLICATION NUMBER: 60/105882
; PRIOR FILING DATE: 1998-10-27
; PRIOR APPLICATION NUMBER: 60/106023
; PRIOR FILING DATE: 1998-10-28

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 81 SSGASTAT 88
|||
Db 39 SSGASTAT 46

RESULT 15

US-10-187-749-310
; Sequence 310, Application US/10187749
; Publication No. US20030153036A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: F3430RIC1
; CURRENT APPLICATION NUMBER: US/10/187,749
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: US/10/052,586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486

/ PRIOR FILING DATE: 1997-10-21
/ PRIOR APPLICATION NUMBER: 60/063540
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063541
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063544
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 310
/ LENGTH: 596
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-187-749-310

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 SSGASTAT 88
Db 39 SSGASTAT 46

RESULT 16

US-10-194-457-310
/ Sequence 310, Application US/10194457
/ Publication No. US20030153037A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P3430R1C296

CURRENT FILING DATE: 2002-07-11

PRIOR FILING DATE: 1997-09-18

PRIOR APPLICATION NUMBER: 60/052586

PRIOR FILING DATE: 2002-01-15

PRIOR APPLICATION NUMBER: 60/059263

PRIOR FILING DATE: 1997-09-18

PRIOR APPLICATION NUMBER: 60/059266

PRIOR FILING DATE: 1997-09-18

PRIOR APPLICATION NUMBER: 60/062250

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/063120

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063121

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063486

PRIOR FILING DATE: 1997-10-21

PRIOR APPLICATION NUMBER: 60/063540

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063541

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 612

SEQ ID NO 310

LENGTH: 596

TYPE: PRT

ORGANISM: Homo Sapien

US-10-194-457-310

; ORGANISM: Homo sapiens
US-10-196-747-310

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
Db 39 SSGASTAT 46

RESULT 19

US-10-015-392A-243

; Sequence 243, Application US/10015392A

; Publication No. US20030166901A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: P2830P1C58

; CURRENT APPLICATION NUMBER: US/10/015,392A

; CURRENT FILING DATE: 2001-12-12

; PRIOR APPLICATION NUMBER: 60/098716

; PRIOR FILING DATE: 1998-09-01

; PRIOR APPLICATION NUMBER: 60/098723

; PRIOR FILING DATE: 1998-09-01

; PRIOR APPLICATION NUMBER: 60/098750

; PRIOR FILING DATE: 1998-09-01

; PRIOR APPLICATION NUMBER: 60/098803

; PRIOR FILING DATE: 1998-09-02

; PRIOR APPLICATION NUMBER: 60/098821

; PRIOR FILING DATE: 1998-09-02

; PRIOR APPLICATION NUMBER: 60/098843

; PRIOR FILING DATE: 1998-09-02

; PRIOR APPLICATION NUMBER: 60/099536

; PRIOR FILING DATE: 1998-09-09

; PRIOR APPLICATION NUMBER: 60/099596

; PRIOR FILING DATE: 1998-09-09

; PRIOR APPLICATION NUMBER: 60/099598

; PRIOR FILING DATE: 1998-09-09

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 477

; SEQ ID NO 243

; LENGTH: 596

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-015-392A-243

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
Db 39 SSGASTAT 46

RESULT 20

US-10-017-253A-243

; Sequence 243, Application US/10017253A

; Publication No. US20030166055A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: P2830P1C62

; CURRENT APPLICATION NUMBER: US/10/017,253A

; CURRENT FILING DATE: 2001-12-13

; PRIOR APPLICATION NUMBER: 60/098716

; PRIOR FILING DATE: 1998-09-01

; PRIOR APPLICATION NUMBER: 60/098723

; PRIOR FILING DATE: 1998-09-01

; PRIOR APPLICATION NUMBER: 60/098749

; PRIOR FILING DATE: 1998-09-01

; PRIOR APPLICATION NUMBER: 60/098750

; PRIOR FILING DATE: 1998-09-01

; PRIOR APPLICATION NUMBER: 60/098803

; PRIOR FILING DATE: 1998-09-02

; PRIOR APPLICATION NUMBER: 60/098821

; PRIOR FILING DATE: 1998-09-02

; PRIOR APPLICATION NUMBER: 60/098843

; PRIOR FILING DATE: 1998-09-02

; PRIOR APPLICATION NUMBER: 60/099536

; PRIOR FILING DATE: 1998-09-09

; PRIOR APPLICATION NUMBER: 60/099596

; PRIOR FILING DATE: 1998-09-09

; PRIOR APPLICATION NUMBER: 60/099598

; PRIOR FILING DATE: 1998-09-09

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 477

; SEQ ID NO 243

; LENGTH: 596

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-017-253A-243

Query Match 2.0%; Score 8; DB 12; Length 596;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88

Db 39 SSGASTAT 46

RESULT 21

US-10-173-689-310

; Sequence 310, Application US/10173689

; Publication No. US20030166104A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnovers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James

```

; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C10
; CURRENT APPLICATION NUMBER: US/10/173,689
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-689-310

Query Match      2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      81 SSGASTAT 88
Db      39 SSGASTAT 46

RESULT 22
US-10-173-690-310
; Sequence 310, Application US/10173690
; Publication No. US20030166105A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C9
; CURRENT APPLICATION NUMBER: US/10/173,690
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-690-310

Query Match      2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      81 SSGASTAT 88
Db      39 SSGASTAT 46

RESULT 23
US-10-173-691-310
; Sequence 310, Application US/10173691
; Publication No. US20030166106A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey

```

```

; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C13
; CURRENT APPLICATION NUMBER: US/10/173,691
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-691-310

```

```

Query Match      2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      81 SSGASTAT 88
Db      39 SSGASTAT 46

```

```

RESULT 24
US-10-173-692-310
; Sequence 310, Application US/10173692
; Publication No. US20030166188A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C20
; CURRENT APPLICATION NUMBER: US/10/173,692
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-692-310

```

```

Query Match      2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      81 SSGASTAT 88
Db      39 SSGASTAT 46

```

```

RESULT 25
US-10-173-694-310
; Sequence 310, Application US/10173694
; Publication No. US20030166107A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

```

```

; APPLICANT: Chen,Jian
; APPLICANT: Desnoyers,Luc
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C19
; CURRENT APPLICATION NUMBER: US/10/173,694
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-694-310

```

```

Query Match      2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      81 SSGASTAT 88
Db      39 SSGASTAT 46

```

RESULT 26

```

US-10-173-698-310
; Sequence 310, Application US/10173698
; Publication No. US20030166108A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C12
; CURRENT APPLICATION NUMBER: US/10/173,698
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-698-310

```

```

Query Match      2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      81 SSGASTAT 88
Db      39 SSGASTAT 46

```

RESULT 27

```

US-10-173-699-310
; Sequence 310, Application US/10173699

```

```

; Publication No. US20030166109A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C8
; CURRENT APPLICATION NUMBER: US/10/173,699
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-699-310

```

```

Query Match      2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      81 SSGASTAT 88
Db      39 SSGASTAT 46

```

RESULT 28

```

US-10-173-707-310
; Sequence 310, Application US/10173707
; Publication No. US20030166110A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C17
; CURRENT APPLICATION NUMBER: US/10/173,707
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-707-310

```

```

Query Match      2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      81 SSGASTAT 88
Db      39 SSGASTAT 46

```

```

RESULT 29
US-10-174-569-310
; Sequence 310, Application US/10174569
; Publication No. US2003016611A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C39
; CURRENT APPLICATION NUMBER: US/10/174,569
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-569-310

Query Match          2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      81 SSGASTAT 88
Db      39 SSGASTAT 46
|||||

RESULT 30
US-10-174-583-310
; Sequence 310, Application US/10174583
; Publication No. US2003016611A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C22
; CURRENT APPLICATION NUMBER: US/10/174,583
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-583-310

Query Match          2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      81 SSGASTAT 88
Db      39 SSGASTAT 46
|||||

RESULT 29
US-10-174-569-310
; Sequence 310, Application US/10174569
; Publication No. US2003016611A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C39
; CURRENT APPLICATION NUMBER: US/10/174,569
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-569-310

Query Match          2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      81 SSGASTAT 88
Db      39 SSGASTAT 46
|||||

RESULT 30
US-10-174-583-310
; Sequence 310, Application US/10174583
; Publication No. US2003016611A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C22
; CURRENT APPLICATION NUMBER: US/10/174,583
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-583-310

Query Match          2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      81 SSGASTAT 88
Db      39 SSGASTAT 46
|||||

```

```

Db      39 SSGASTAT 46

RESULT 31
US-10-174-587-310
; Sequence 310, Application US/10174587
; Publication No. US2003016611A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C30
; CURRENT APPLICATION NUMBER: US/10/174,587
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-587-310

Query Match          2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      81 SSGASTAT 88
Db      39 SSGASTAT 46
|||||

RESULT 32
US-10-174-589-310
; Sequence 310, Application US/10174589
; Publication No. US2003016611A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C33
; CURRENT APPLICATION NUMBER: US/10/174,589
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-589-310

Query Match          2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 81 SSGASTAT 88
| | | | | | | |
Db 39 SSGASTAT 46

RESULT 33

US-10-174-591-310
; Sequence 310, Application US/10174591
; Publication No. US20030166115A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C32
; CURRENT APPLICATION NUMBER: US/10/174,591
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-591-310

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
| | | | | | | |
Db 39 SSGASTAT 46

RESULT 34

US-10-175-736-310
; Sequence 310, Application US/10175736
; Publication No. US20030166117A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C43
; CURRENT APPLICATION NUMBER: US/10/175,736
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-736-310

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
| | | | | | | |
Db 39 SSGASTAT 46

RESULT 35

US-10-175-742-310
; Sequence 310, Application US/10175742
; Publication No. US20030166118A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C59
; CURRENT APPLICATION NUMBER: US/10/175,742
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-742-310

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
| | | | | | | |
Db 39 SSGASTAT 46

RESULT 36

US-10-175-744-310
; Sequence 310, Application US/10175744
; Publication No. US20030166119A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C49
; CURRENT APPLICATION NUMBER: US/10/175,744
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
US-10-175-744-310

; ORGANISM: Homo Sapien
US-10-175-744-310

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
Db 39 SSGASTAT 46

RESULT 37

US-10-175-745-310
; Sequence 310, Application US/10175745
; Publication No. US20030166120A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C55
; CURRENT APPLICATION NUMBER: US/10/175,745
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-745-310

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
Db 39 SSGASTAT 46

RESULT 38

US-10-175-748-310
; Sequence 310, Application US/10175748
; Publication No. US20030166121A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C62
; CURRENT APPLICATION NUMBER: US/10/175,748
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-748-310

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
Db 39 SSGASTAT 46

RESULT 39

US-10-175-751-310
; Sequence 310, Application US/10175751
; Publication No. US20030166122A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C57
; CURRENT APPLICATION NUMBER: US/10/175,751
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-751-310

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
Db 39 SSGASTAT 46

RESULT 40

US-10-175-754-310
; Sequence 310, Application US/10175754
; Publication No. US20030166123A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C58
; CURRENT APPLICATION NUMBER: US/10/175,754

; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-754-310

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 SSGASTAT 88
| | | | |
Db 39 SSGASTAT 46

RESULT 41

US-10-176-480-310
; Sequence 310, Application US/10176480
; Publication No. US20030166124A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C96
; CURRENT APPLICATION NUMBER: US/10/176,480
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-480-310

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 SSGASTAT 88
| | | | |
Db 39 SSGASTAT 46

RESULT 42

US-10-176-489-310
; Sequence 310, Application US/10176489
; Publication No. US20030166125A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C102
; CURRENT APPLICATION NUMBER: US/10/176,489
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-489-310

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 SSGASTAT 88
| | | | |
Db 39 SSGASTAT 46

RESULT 43

US-10-176-754-310
; Sequence 310, Application US/10176754
; Publication No. US20030166126A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C85
; CURRENT APPLICATION NUMBER: US/10/176,754
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-754-310

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 SSGASTAT 88
| | | | |
Db 39 SSGASTAT 46

RESULT 44

US-10-176-755-310
; Sequence 310, Application US/10176755
; Publication No. US20030166127A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.


```

; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C106
; CURRENT APPLICATION NUMBER: US/10/176,755
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-755-310

Query Match      2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      81 SSGASTAT 88
Db      39 SSGASTAT 46

RESULT 45
US-10-176-759-310
; Sequence 310, Application US/10176759
; Publication No. US20030166128A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C80
; CURRENT APPLICATION NUMBER: US/10/176,759
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-759-310

Query Match      2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      81 SSGASTAT 88
Db      39 SSGASTAT 46

RESULT 46
US-10-176-920-310
; Sequence 310, Application US/10176920
; Publication No. US20030166129A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C120
; CURRENT APPLICATION NUMBER: US/10/176,922
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-920-310

Query Match      2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      81 SSGASTAT 88
Db      39 SSGASTAT 46

RESULT 47
US-10-176-922-310
; Sequence 310, Application US/10176922
; Publication No. US20030166130A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C120
; CURRENT APPLICATION NUMBER: US/10/176,922
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-922-310

Query Match      2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      81 SSGASTAT 88
Db      39 SSGASTAT 46

RESULT 48
US-10-176-924-310
; Sequence 310, Application US/10176924
; Publication No. US20030166131A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc

```



```
; Sequence 310, Application US/10179515
; Publication No. US20030166135A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C131
; CURRENT APPLICATION NUMBER: US/10/179,515
; CURRENT FILING DATE: 2002-06-24
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-179-515-310
```

```
Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 81 SSGASTAT 88
Db 39 SSGASTAT 46
```

RESULT 53

```
US-10-017-306A-243
; Sequence 243, Application US/10017306A
; Publication No. US20030170718A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, Nicholas
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830PIC66
; CURRENT APPLICATION NUMBER: US/10/017,306A
; CURRENT FILING DATE: 2002-06-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 243
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-306A-243
```

```
Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 81 SSGASTAT 88
```

```
Db 39 SSGASTAT 46
```

RESULT 54

```
US-10-063-526-100
; Sequence 100, Application US/10063526
; Publication No. US20030171550A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,526
; CURRENT FILING DATE: 2002-05-02
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 100
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-526-100
```

```
Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 81 SSGASTAT 88
Db 39 SSGASTAT 46
```

RESULT 55

```
US-10-173-702-310
; Sequence 310, Application US/10173702
; Publication No. US20030170793A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C15
; CURRENT APPLICATION NUMBER: US/10/173,702
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-702-310
```

```
Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 81 SSGASTAT 88
 |||||
 Db 39 SSGASTAT 46

RESULT 56

US-10-173-703-310
 ; Sequence 310, Application US/10173703
 ; Publication No. US20030170794A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3430R1C16
 ; CURRENT FILING DATE: 2002-06-17
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 310
 ; LENGTH: 596
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-173-703-310

Query Match 2.0%; Score 8; DB 12; Length 596;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 SSGASTAT 88
 |||||
 Db 39 SSGASTAT 46

RESULT 57

US-10-173-704-310
 ; Sequence 310, Application US/10173704
 ; Publication No. US20030170795A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3430R1C2
 ; CURRENT FILING DATE: 2002-06-17
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 310
 ; LENGTH: 596
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-173-704-310

Query Match 2.0%; Score 8; DB 12; Length 596;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 SSGASTAT 88
 |||||
 Db 39 SSGASTAT 46

RESULT 58

US-10-174-574-310
 ; Sequence 310, Application US/10174574
 ; Publication No. US20030170796A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3430R1C25
 ; CURRENT APPLICATION NUMBER: US/10/174,574
 ; CURRENT FILING DATE: 2002-06-18
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 310
 ; LENGTH: 596
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-174-574-310

Query Match 2.0%; Score 8; DB 12; Length 596;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 SSGASTAT 88
 |||||
 Db 39 SSGASTAT 46

RESULT 59

US-10-176-486-310
 ; Sequence 310, Application US/10176486
 ; Publication No. US20030170797A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3430R1C13
 ; CURRENT APPLICATION NUMBER: US/10/176,486
 ; CURRENT FILING DATE: 2002-06-21
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 310
 ; LENGTH: 596
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien

US-10-176-486-310

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
Db 39 SSGASTAT 46

RESULT 60

US-10-176-490-310
; Sequence 310, Application US/10176490
; Publication No. US20030170798A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C79
; CURRENT APPLICATION NUMBER: US/10/176,490
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-490-310

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
Db 39 SSGASTAT 46

RESULT 61

US-10-176-752-310
; Sequence 310, Application US/10176752
; Publication No. US20030170799A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C101
; CURRENT APPLICATION NUMBER: US/10/176,752
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310

; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-752-310

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
Db 39 SSGASTAT 46

RESULT 62

US-10-176-981-310
; Sequence 310, Application US/10176981
; Publication No. US20030170800A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C73
; CURRENT APPLICATION NUMBER: US/10/176,981
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-981-310

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
Db 39 SSGASTAT 46

RESULT 63

US-10-176-983-310
; Sequence 310, Application US/10176983
; Publication No. US20030170801A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C97
; CURRENT APPLICATION NUMBER: US/10/176,983
; CURRENT FILING DATE: 2002-06-21

; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-983-310

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
Db 39 SSGASTAT 46

RESULT 64

US-10-176-988-310
; Sequence 310, Application US/10176988
; Publication No. US20030170802A1

; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: F3430R1C87

; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien

US-10-176-988-310

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
Db 39 SSGASTAT 46

RESULT 65

US-10-179-517-310

; Sequence 310, Application US/10179517
; Publication No. US20030170805A1

; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: F3430R1C87

; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien

; FILE REFERENCE: P3430R1C136
; CURRENT APPLICATION NUMBER: US/10/179,517
; CURRENT FILING DATE: 2002-06-24
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-179-517-310

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
Db 39 SSGASTAT 46

RESULT 66

US-10-179-521-310

; Sequence 310, Application US/10179521
; Publication No. US20030170806A1

; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C139

; CURRENT FILING DATE: 2002-06-24
; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien

US-10-179-521-310

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
Db 39 SSGASTAT 46

RESULT 67

US-10-063-586-100

; Sequence 100, Application US/10063586
; Publication No. US20030176684A1

; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.

; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063.586
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 100
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-586-100

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 SSGASTAT 88
Db 39 SSGASTAT 46

RESULT 68

US-10-012-064A-243
; Sequence 243, Application US/10012064A
; Publication No. US20030180836A1

; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan I.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, Christopher J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas P.

; TITLE OF INVENTION: Secured and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: P2830P1C19

; CURRENT APPLICATION NUMBER: US/10/012.064A

; CURRENT FILING DATE: 2002-07-15

; PRIOR APPLICATION NUMBER: 60/098716

; PRIOR FILING DATE: 1998-09-01

; PRIOR APPLICATION NUMBER: 60/098723

; PRIOR FILING DATE: 1998-09-01

; PRIOR APPLICATION NUMBER: 60/098749

; PRIOR FILING DATE: 1998-09-01

; PRIOR APPLICATION NUMBER: 60/098750

; PRIOR FILING DATE: 1998-09-01

; PRIOR APPLICATION NUMBER: 60/098803

; PRIOR FILING DATE: 1998-09-02

; PRIOR APPLICATION NUMBER: 60/098821

; PRIOR FILING DATE: 1998-09-02

; PRIOR APPLICATION NUMBER: 60/098843

; PRIOR FILING DATE: 1998-09-02

; PRIOR APPLICATION NUMBER: 60/099536

; PRIOR FILING DATE: 1998-09-09

; PRIOR APPLICATION NUMBER: 60/099596

; PRIOR FILING DATE: 1998-09-09

; PRIOR APPLICATION NUMBER: 60/099598

; PRIOR FILING DATE: 1998-09-09

; Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 477

; SEQ ID NO 243

; LENGTH: 596

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-012-064A-243

Query Match 2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 SSGASTAT 88
Db 39 SSGASTAT 46

RESULT 69

US-10-017-867A-243

; Sequence 243, Application US/10017867A

; Publication No. US20030180792A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan I.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, Christopher J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas P.

; TITLE OF INVENTION: Secured and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: P2830P1C60

; CURRENT APPLICATION NUMBER: US/10/017.867A

; CURRENT FILING DATE: 2001-12-13

; PRIOR APPLICATION NUMBER: 60/098716

; PRIOR FILING DATE: 1998-09-01

; PRIOR APPLICATION NUMBER: 60/098723

; PRIOR FILING DATE: 1998-09-01

; PRIOR APPLICATION NUMBER: 60/098749

; PRIOR FILING DATE: 1998-09-01

; PRIOR APPLICATION NUMBER: 60/098750

; PRIOR FILING DATE: 1998-09-01

; PRIOR APPLICATION NUMBER: 60/098803

; PRIOR FILING DATE: 1998-09-02

; PRIOR APPLICATION NUMBER: 60/098821

; PRIOR FILING DATE: 1998-09-02

; PRIOR APPLICATION NUMBER: 60/098843

; PRIOR FILING DATE: 1998-09-02

; PRIOR APPLICATION NUMBER: 60/099536

; PRIOR FILING DATE: 1998-09-09

; PRIOR APPLICATION NUMBER: 60/099596

; PRIOR FILING DATE: 1998-09-09

; PRIOR APPLICATION NUMBER: 60/099598

; PRIOR FILING DATE: 1998-09-09

; PRIOR APPLICATION NUMBER: 60/099602

; PRIOR FILING DATE: 1998-09-09

; PRIOR APPLICATION NUMBER: 60/099642

; PRIOR FILING DATE: 1998-09-09

; PRIOR APPLICATION NUMBER: 60/099741

; PRIOR FILING DATE: 1998-09-10

; PRIOR APPLICATION NUMBER: 60/099754

; PRIOR FILING DATE: 1998-09-10

; PRIOR APPLICATION NUMBER: 60/099763

; PRIOR FILING DATE: 1998-09-10

; PRIOR APPLICATION NUMBER: 60/099792

; PRIOR FILING DATE: 1998-09-10

; PRIOR APPLICATION NUMBER: 60/099808

; PRIOR FILING DATE: 1998-09-10

; PRIOR APPLICATION NUMBER: 60/099812

; PRIOR FILING DATE: 1998-09-10

; PRIOR APPLICATION NUMBER: 60/099815

; PRIOR FILING DATE: 1998-09-10

; PRIOR APPLICATION NUMBER: 60/099816

; PRIOR FILING DATE: 1998-09-10

;
; PRIOR APPLICATION NUMBER: 60/100385
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100388
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100390
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100584
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100627
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100661
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100662
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100664
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100683
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100684
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100710
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100711
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100848
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/100849
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/100919
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100930
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/101014
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/101068
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/101071
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/101279
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: 60/101471
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101472
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101474
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101475
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101476
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101477
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101479
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101738
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101741
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101743
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101915
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101916
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/102207
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/102240
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/102307
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/102330
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/102331

;
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/102484
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 60/102487
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 60/102570
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 60/102571
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 60/102684
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 60/102687
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 60/102965
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 60/103258
; PRIOR FILING DATE: 1998-10-06
; PRIOR APPLICATION NUMBER: 60/103314
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103315
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103328
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103395
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103396
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103401
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103449
; PRIOR FILING DATE: 1998-10-06
; PRIOR APPLICATION NUMBER: 60/103633
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/103678
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/103679
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/103711
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/104257
; PRIOR FILING DATE: 1998-10-14
; PRIOR APPLICATION NUMBER: 60/104987
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105000
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105002
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105104
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 60/105169
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: 60/105266
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: 60/105693
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: 60/105694
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: 60/105807
; PRIOR FILING DATE: 1998-10-27
; PRIOR APPLICATION NUMBER: 60/105881
; PRIOR FILING DATE: 1998-10-27
; PRIOR APPLICATION NUMBER: 60/105882
; PRIOR FILING DATE: 1998-10-27
; PRIOR APPLICATION NUMBER: 60/106023
; PRIOR FILING DATE: 1998-10-28

Query Match 2.0%; Score 8; DB 12; Length 596;

Best Local Similarity 100.0%; Pred. No. 1.2e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0;

QY 81 SSGASTAT 88

Db 39 SSGASTAT 46

RESULT 70

US-10-063-510-100
 ; Sequence 100, Application US/10063510
 ; Publication No. US20030180837A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3230R1C1
 ; CURRENT APPLICATION NUMBER: US/10/063,510
 ; Prior Filing Date: 2002-05-01
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 170
 ; SEQ ID NO 100
 ; LENGTH: 596
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-063-510-100

Query Match 2.0%; Score 8; DB 12; Length 596;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
 DB 39 SSGASTAT 46

RESULT 71

US-10-063-514-100
 ; Sequence 100, Application US/10063514
 ; Publication No. US20030181707A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3230R1C1
 ; CURRENT APPLICATION NUMBER: US/10/063,514
 ; Prior Filing Date: 2002-05-01
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 170
 ; SEQ ID NO 100
 ; LENGTH: 596
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-063-514-100

Query Match 2.0%; Score 8; DB 12; Length 596;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
 DB 39 SSGASTAT 46

RESULT 72

US-10-063-516-100
 ; Sequence 100, Application US/10063516
 ; Publication No. US20030181708A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3230R1C1
 ; CURRENT APPLICATION NUMBER: US/10/063,516
 ; Prior Filing Date: 2002-05-01
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 170
 ; SEQ ID NO 100
 ; LENGTH: 596
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-063-516-100

Query Match 2.0%; Score 8; DB 12; Length 596;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
 DB 39 SSGASTAT 46

RESULT 73

US-10-063-523-100
 ; Sequence 100, Application US/10063523
 ; Publication No. US20030181636A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3230R1C1
 ; CURRENT APPLICATION NUMBER: US/10/063,523
 ; Prior Filing Date: 2002-05-02
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 170
 ; SEQ ID NO 100
 ; LENGTH: 596
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-063-523-100

Query Match 2.0%; Score 8; DB 12; Length 596;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SSGASTAT 88
 DB 39 SSGASTAT 46

Job time : 36.0556 secs

```

RESULT 74
US-10-063-527-100
; Sequence 100, Application US/10063527
; Publication No. US20030181637A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,527
; CURRENT FILING DATE: 2002-05-02
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 100
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-527-100

Query Match      2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      81 SSGASTAT 88
Db      39 SSGASTAT 46

```

```

RESULT 75
US-10-063-528-100
; Sequence 100, Application US/10063528
; Publication No. US20030181666A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,528
; CURRENT FILING DATE: 2002-05-02
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 100
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-528-100

Query Match      2.0%; Score 8; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      81 SSGASTAT 88
Db      39 SSGASTAT 46

```

Search completed: December 17, 2003, 06:42:59

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 06:32:02 ; Search time 14.1462 Seconds
(without alignments)
1211.344 Million cell updates/sec

Title: US-09-805-550-2

Perfect score: 405

Sequence: 1 MKLVNKLKGTNFEIASPD.....EELTANYLLDHGHEFDQQQ 405

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	405	100.0	405	3	US-09-413-574-2
2	13	3.2	368	3	US-09-413-574-4
3	8	2.0	23	2	US-08-492-027A-3
4	8	2.0	117	6	5242821-12
5	8	2.0	177	4	US-09-252-991A-28452
6	8	2.0	180	4	US-09-252-991A-18890
7	8	2.0	212	4	US-09-252-991A-24512
8	8	2.0	272	2	US-08-492-027A-4
9	8	2.0	495	4	US-09-252-991A-31949
10	8	2.0	568	4	US-09-252-991A-23264
11	8	2.0	720	4	US-09-252-991A-31915
12	8	2.0	892	4	US-09-585-858-16
13	8	2.0	898	2	US-08-808-982-5
14	8	2.0	898	3	US-09-306-902A-5
15	7	1.7	62	1	US-08-428-091-1
16	7	1.7	78	1	US-07-881-075-5
17	7	1.7	78	1	US-08-120-827-5
18	7	1.7	78	1	US-08-478-675-5
19	7	1.7	85	1	US-08-428-091-4
20	7	1.7	104	4	US-09-134-001C-4716
21	7	1.7	150	4	US-09-252-991A-30419
22	7	1.7	155	4	US-09-252-991A-20273
23	7	1.7	159	4	US-09-252-991A-18088
24	7	1.7	179	4	US-09-252-991A-26223
25	7	1.7	191	3	US-08-974-022-52
26	7	1.7	191	3	US-08-795-445A-52
27	7	1.7	191	3	US-08-795-447A-52

Sequence 52, Appl	191	3	US-08-974-186-52	1.7
Sequence 52, Appl	191	3	US-08-795-446B-52	1.7
Sequence 139, App	191	4	US-08-706-945D-139	1.7
Sequence 155, App	221	4	US-09-904-615-155	1.7
Sequence 6, Appl	256	1	US-08-236-918A-6	1.7
Sequence 2, Appl	256	4	US-09-150-864A-6	1.7
Sequence 2, Appl	256	4	US-08-012-269A-2	1.7
Sequence 2, Appl	256	5	PCT-US96-03965-2	1.7
Sequence 2764, A	258	4	US-09-252-991A-37264	1.7
Sequence 17461, A	272	4	US-09-252-991A-17461	1.7
Sequence 17461, A	281	1	US-08-397-633A-75	1.7
Sequence 26761, A	281	4	US-09-252-991A-26761	1.7
Sequence 76, Appl	287	1	US-08-397-633A-76	1.7
Sequence 22293, A	296	4	US-09-252-991A-22293	1.7
Sequence 17, Appl	303	4	US-09-495-406-17	1.7
Sequence 3568, Ap	341	4	US-09-134-001C-3568	1.7
Sequence 24458, A	343	4	US-09-252-991A-24458	1.7
Sequence 6666, Ap	360	4	US-09-107-532A-6666	1.7
Sequence 26295, A	364	4	US-09-252-991A-26295	1.7
Sequence 23727, A	382	4	US-09-252-991A-23727	1.7
Sequence 18, Appl	391	3	US-09-377-557-18	1.7
Sequence 21075, A	406	4	US-09-252-991A-21075	1.7
Sequence 27669, A	415	4	US-09-252-991A-27669	1.7
Sequence 22253, A	420	4	US-09-252-991A-22253	1.7
Sequence 23618, A	423	4	US-09-252-991A-23618	1.7
Sequence 2, Appl	448	3	US-08-476-509B-2	1.7
Sequence 6, Appl	465	3	US-08-948-997-6	1.7
Sequence 6, Appl	465	3	US-09-348-817A-6	1.7
Sequence 6, Appl	465	4	US-09-722-292-6	1.7
Sequence 23144, A	479	4	US-09-252-991A-23144	1.7
Sequence 2, Appl	486	3	US-08-348-518C-2	1.7
Sequence 27501, A	495	4	US-09-252-991A-27501	1.7
Sequence 16737, A	519	4	US-09-252-991A-16737	1.7
Sequence 22537, A	534	4	US-09-252-991A-22537	1.7
Sequence 26327, A	549	4	US-09-252-991A-26327	1.7
Sequence 23080, A	561	4	US-09-252-991A-23080	1.7
Sequence 573, App	567	4	US-09-205-258-573	1.7
Sequence 7, Appl	591	3	US-09-199-290-7	1.7
Sequence 23472, A	617	4	US-09-252-991A-23472	1.7
Sequence 34, Appl	618	3	US-09-199-290-34	1.7
Sequence 21585, A	619	4	US-09-252-991A-21585	1.7
Sequence 28001, A	657	4	US-09-252-991A-28001	1.7
Sequence 14, Appl	661	2	US-08-795-868-14	1.7
Sequence 14, Appl	661	4	US-09-303-069-14	1.7
Sequence 14, Appl	661	4	US-09-134-250-14	1.7
Sequence 20752, A	749	4	US-09-252-991A-20752	1.7
Sequence 23774, A	798	4	US-09-252-991A-23774	1.7
Sequence 25050, A	802	4	US-09-252-991A-25050	1.7
Sequence 2, Appl	803	3	US-09-118-442-2	1.7
Sequence 2, Appl	803	3	US-09-677-064-2	1.7
Sequence 25143, A	848	4	US-09-252-991A-25143	1.7
Sequence 3, Appl	872	1	US-08-491-357-3	1.7
Sequence 3, Appl	872	3	US-08-968-633-3	1.7
Sequence 3, Appl	872	3	US-09-196-466-3	1.7
Sequence 3, Appl	872	5	PCT-US96-10823-3	1.7
Sequence 54, Appl	877	1	US-08-397-633A-54	1.7
Sequence 31458, A	906	4	US-09-252-991A-31458	1.7
Sequence 32183, A	952	4	US-09-252-991A-32183	1.7
Sequence 32419, A	1004	3	US-08-916-352-2	1.7
Sequence 32419, A	1011	4	US-09-252-991A-32419	1.7
Sequence 32, Appl	1042	4	US-09-512-250C-32	1.7
Sequence 2, Appl	1049	2	US-08-817-090B-2	1.7
Sequence 4, Appl	1050	2	US-08-817-090B-4	1.7
Sequence 19328, A	1096	4	US-09-252-991A-19328	1.7
Sequence 14, Appl	1101	2	US-08-916-917-14	1.7
Sequence 14, Appl	1101	3	US-09-225-170-14	1.7
Sequence 4, Appl	1102	2	US-08-916-917-4	1.7
Sequence 4, Appl	1102	2	US-08-972-631-4	1.7
Sequence 4, Appl	1102	2	US-08-972-630-4	1.7
Sequence 4, Appl	1102	2	US-08-972-630-4	1.7
Sequence 4, Appl	1102	2	US-08-672-211-4	1.7
Sequence 4, Appl	1102	3	US-09-225-170-4	1.7
Sequence 23140, A	1246	4	US-09-252-991A-23140	1.7

ALIGNMENTS

RESULT 1

US-09-413-574-2
; Sequence 2, Application US/09413574
; Patent No. 6235972
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; APPLICANT: Taghiani, Laura
; TITLE OF INVENTION: Maize Rad23 Genes and Uses Thereof
; FILE REFERENCE: 0964
; CURRENT APPLICATION NUMBER: US/09/413,574
; CURRENT FILING DATE: 1999-10-06
; EARLIER APPLICATION NUMBER: 60/109,728
; EARLIER FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Zea mays
US-09-413-574-2

Query Match 100.0%; Score 405; DB 3; Length 405;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MKLVNVTLLKGNFIEASPDASVADVKRIETTTQGSTYRADQQMLIYQGKILKDETTL	60
Db	1	MKLVNVTLLKGNFIEASPDASVADVKRIETTTQGSTYRADQQMLIYQGKILKDETTL	60
Qy	61	SNGVAENSVFLVIMLSKAKASSGASTATTAKAPATLAQPAAPVAPAAASVARTPTQAPVAT	120
Db	61	SNGVAENSVFLVIMLSKAKASSGASTATTAKAPATLAQPAAPVAPAAASVARTPTQAPVAT	120
Qy	121	AETAPPSVQQAAPAAATVAATDDADVTYSQASNLVFGNNLEBQTTQQILDMGGGTWERDVT	180
Db	121	AETAPPSVQQAAPAAATVAATDDADVTYSQASNLVFGNNLEBQTTQQILDMGGGTWERDVT	180
Qy	181	VRALRAAYNNPERAIDLYSGIPENVEAQPVARAPAAAGQQTNOQAASPAQPAVALPVQPS	240
Db	181	VRALRAAYNNPERAIDLYSGIPENVEAQPVARAPAAAGQQTNOQAASPAQPAVALPVQPS	240
Qy	241	PASAGPNANPLNLPFGVPSGSGNPGVVPVPGAGSGALDALROLPOFQALLQVLQVQNPQILQ	300
Db	241	PASAGPNANPLNLPFGVPSGSGNPGVVPVPGAGSGALDALROLPOFQALLQVLQVQNPQILQ	300
Qy	301	PMQLGKQNPQIIRLQENQAEFLRLVNESPEGGPGGNILGQLAAAVPQTLTVTTPERE	360
Db	301	PMQLGKQNPQIIRLQENQAEFLRLVNESPEGGPGGNILGQLAAAVPQTLTVTTPERE	360
Qy	361	AIQRLGKGFNRELVLVFFACNKDEELTANYLLDHGHEFDQDQ 405	
Db	361	AIQRLGKGFNRELVLVFFACNKDEELTANYLLDHGHEFDQDQ 405	

RESULT 2

US-09-413-574-4
; Sequence 4, Application US/09413574
; Patent No. 6235972
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; APPLICANT: Taghiani, Laura
; TITLE OF INVENTION: Maize Rad23 Genes and Uses Thereof
; FILE REFERENCE: 0964
; CURRENT APPLICATION NUMBER: US/09/413,574
; CURRENT FILING DATE: 1999-10-06
; EARLIER APPLICATION NUMBER: 60/109,728
; EARLIER FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Zea mays
US-09-413-574-4

Query Match 3.2%; Score 13; DB 3; Length 368;
Best Local Similarity 100.0%; Pred. No. 0.00061;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	182	RALRAAYNNPERA	194
Db	167	RALRAAYNNPERA	179

RESULT 3

US-08-492-027A-3
; Sequence 3, Application US/08492027A
; Patent No. 5912333
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shoichi
; APPLICANT: Burnell, James N
; TITLE OF INVENTION: DNA Encoding Carbonic Anhydrase
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch and Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/492,027A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr, Gerald M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 0760-206
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-492-027A-3

Query Match 2.0%; Score 8; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	100	AAPVAPAA	107
Db	1	AAPVAPAA	8

RESULT 4

5242821-12
; Patent No. 5242821
; APPLICANT: PALVA, LIKKA;SIBAKOV, MERV
; TITLE OF INVENTION: LACTOCOCCUS PROMOTER AND SIGNAL
; SEQUENCES FOR EXPRESSION IN BACTERIA
; NUMBER OF SEQUENCES: 27

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/377,450
; FILING DATE: 10-JUL-1989
; SEQ ID NO:12:
; LENGTH: 117
5242821-12

Query Match 2.0%; Score 8; DB 6; Length 117;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 SKAKASS 82
Db 72 SKAKASS 79

RESULT 5

US-09-252-991A-28452
; Sequence 28452, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28452
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28452

Query Match 2.0%; Score 8; DB 4; Length 177;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 APVAPAS 108
Db 36 APVAPAS 43

RESULT 6

US-09-252-991A-18890
; Sequence 18890, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18890
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18890

Query Match 2.0%; Score 8; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 APAATVAA 140
Db 115 APAATVAA 122

RESULT 7

US-09-252-991A-24512
; Sequence 24512, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24512
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (43)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-24512

Query Match 2.0%; Score 8; DB 4; Length 212;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 PAAPVAPA 106
Db 140 PAAPVAPA 147

RESULT 8

US-08-492-027A-4
; Sequence 4, Application US/08492027A
; Patent No. 5912333
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shoichi
; TITLE OF INVENTION: DNA Encoding Carbonic Anhydrase
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch and Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/492,027A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr, Gerald M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 0760-206
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 272 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-492-027A-4

Query Match 2.0%; Score 8; DB 2; Length 272;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 AAPVAPAA 107
Db 63 AAPVAPAA 70
|||||

RESULT 9

US-09-252-991A-31949
; Sequence 31949, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 31949

; LENGTH: 495

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-31949

Query Match 2.0%; Score 8; DB 4; Length 495;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 212 ARAPAAQ 219
Db 131 ARAPAAQ 138
|||||

RESULT 10

US-09-252-991A-23264

; Sequence 23264, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 23264

; LENGTH: 568

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-23264

Query Match 2.0%; Score 8; DB 4; Length 568;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 AQAAPVA 104
Db 257 AQAAPVA 264
|||||

RESULT 11

US-09-252-991A-31915

; Sequence 31915, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 31915

; LENGTH: 720

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-31915

Query Match 2.0%; Score 8; DB 4; Length 720;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 AAPVAPAA 107
Db 465 AAPVAPAA 472
|||||

RESULT 12

US-09-585-858-16

; Sequence 16, Application US/09585858

; Patent No. 6492161

; GENERAL INFORMATION:

; APPLICANT: Sigridur Hjorleifsdottir

; APPLICANT: Gudmundur O. Hreggvidsson

; APPLICANT: Olafur H. Fridjonsson

; APPLICANT: Arnthor Aevarsson

; APPLICANT: Jakob K. Kristjansson

; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic

; FILE OF INVENTION: Host Organism

; FILE REFERENCE: 2739.1001-001

; CURRENT APPLICATION NUMBER: US/09/585,858

; CURRENT FILING DATE: 2000-12-18

; PRIOR APPLICATION NUMBER: 60/137,120

; PRIOR FILING DATE: 1999-06-02

; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 16

; LENGTH: 892

; TYPE: PRT

; ORGANISM: Herpes simplex virus (type 1/strain 17)

US-09-585-858-16

Query Match 2.0%; Score 8; DB 4; Length 892;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 TLAQPAAP 102
Db 38 TLAQPAAP 45
|||||

RESULT 13

US-08-808-982-5

; Sequence 5, Application US/08808982

```
; Patent No. 5939271
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Leonardo, E. David
; APPLICANT: Hink, Lindsey
; APPLICANT: Masu, Masayuki
; APPLICANT: Kazuko, Keino-Masu
; NUMBER OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,982
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 898 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; US-08-808-982-5

Query Match          2.0%; Score 8; DB 2; Length 898;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      341 LGQLAAAV 348
Db      873 LGQLAAAV 880

RESULT 14
US-09-306-902A-5
; Sequence 5, Application US/09306902A
; Patent No. 627585
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Leonardo, E. David
; APPLICANT: Hink, Lindsey
; APPLICANT: Masu, Masayuki
; APPLICANT: Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/306,902A
; FILING DATE: 07-May-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 898 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; US-09-306-902A-5

Query Match          2.0%; Score 8; DB 3; Length 898;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      341 LGQLAAAV 348
Db      873 LGQLAAAV 880

RESULT 15
US-08-428-091-1
; Sequence 1, Application US/08428091
; Patent No. 568380
; GENERAL INFORMATION:
; APPLICANT: GERMOND, JACQUES.-EDOUARD
; APPLICANT: MARCISSET, OLIVIER
; APPLICANT: MOLLET, BEAT
; TITLE OF INVENTION: BACTERIOICINS OF STREPTOCOCCUS
; TITLE OF INVENTION: THERMOPHILUS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 AVENUE OF THE AMERICA
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/428,091
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP PCT/EP94/02805
; FILING DATE: 24-AUG-1994
; APPLICATION NUMBER: CH 2628/93-7
; FILING DATE: 03-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: FANUCCI, ALLAN A
; REGISTRATION NUMBER: 30256
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790 9090
; TELEFAX: 212 869 8864
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 62 amino acids
; TYPE: amino acid
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus thermophilus
; STRAIN: CNM I-1351
US-08-428-091-1

Query Match      1.7%  Score 7;  DB 1;  Length 62;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      270 GAGSGAL 276
Db      17 GAGSGAL 23

RESULT 16
US-07-881-075-5
; Sequence 5, Application US/07881075
; Patent No. 5444149
; GENERAL INFORMATION:
; APPLICANT: KEENE, JACK D.
; APPLICANT: KING, PETER H.
; APPLICANT: LEVINE, TODD
; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL IN THE
; TITLE OF INVENTION: RECOGNITION, BINDING AND EXPRESSION OF RIBONUCLEIC ACIDS
; TITLE OF INVENTION: INVOLVED IN CELL GROWTH, NEOPLASIA AND IMMUNOREGULATION
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 19920511
; APPLICATION NUMBER: US/07/881,075
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5444149man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 714-154-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-07-881-075-5

Query Match      1.7%  Score 7;  DB 1;  Length 78;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 7;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      284 QFQALLQ 290
Db      37 QFQALLQ 43

RESULT 17
US-08-120-827-5
; Sequence 5, Application US/08120827
```

```
; Patent No. 5525495
; GENERAL INFORMATION:
; APPLICANT: KEENE, JACK D.
; APPLICANT: KING, PETER H.
; APPLICANT: LEVINE, TODD
; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL IN THE
; TITLE OF INVENTION: RECOGNITION, BINDING AND EXPRESSION OF RIBONUCLEIC ACIDS
; TITLE OF INVENTION: INVOLVED IN CELL GROWTH, NEOPLASIA AND IMMUNOREGULATION
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 15-SEP-1993
; APPLICATION NUMBER: US/08/120,827
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5525495man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 714-158-0 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)413-3000
; TELEFAX: (703)413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-120-827-5

Query Match      1.7%  Score 7;  DB 1;  Length 78;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 7;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      284 QFQALLQ 290
Db      37 QFQALLQ 43

RESULT 18
US-08-478-675-5
; Sequence 5, Application US/08478675
; Patent No. 5773246
; GENERAL INFORMATION:
; APPLICANT: KEENE, JACK D.
; APPLICANT: KING, PETER H.
; APPLICANT: LEVINE, TODD
; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL IN THE
; TITLE OF INVENTION: RECOGNITION, BINDING AND EXPRESSION OF RIBONUCLEIC ACIDS
; TITLE OF INVENTION: INVOLVED IN CELL GROWTH, NEOPLASIA AND IMMUNOREGULATION
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```



```
;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,675
; FILING DATE: 07-JUN-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/120,827
; FILING DATE: 15-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Orlon, No. 577346man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 714-158-0 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)413-3000
; TELEFAX: (703)413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-478-675-5

Query Match      1.7%; Score 7; DB 1; Length 78;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      284 QFOALLQ 290
DB      37 QFOALLQ 43

RESULT 19
US-08-428-091-4
; Sequence 4, Application US/08428091
; Patent No. 5681890
; GENERAL INFORMATION:
; APPLICANT: GERMOND, JACQUES.-EDOUARD
; APPLICANT: MARCISSET, OLIVIER
; APPLICANT: MOLLET, BEAT
; TITLE OF INVENTION: BACTERIOCINS OF STREPTOCOCCUS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 AVENUE OF THE AMERICA
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/428,091
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP PCT/EP94/02805
; FILING DATE: 24-AUG-1994
; APPLICATION NUMBER: CH 2628/93-7
; FILING DATE: 03-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: FANUCCI, ALLAN A
; REGISTRATION NUMBER: 30256
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790 9090
; TELEFAX: 212 869 8864
```

```
;
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 85 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-428-091-4
```

```
Query Match      1.7%; Score 7; DB 1; Length 85;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      270 GAGSGAL 276
DB      40 GAGSGAL 46
```

RESULT 20

```
US-09-134-001C-4716
; Sequence 4716, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4716
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-4716
```

```
Query Match      1.7%; Score 7; DB 4; Length 104;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      82 SGASTAT 88
DB      70 SGASTAT 76
```

RESULT 21

```
US-09-252-991A-30419
; Sequence 30419, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30419
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-30419
```

```
Query Match      1.7%; Score 7; DB 4; Length 150;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 97 AQAAPV 103
|||||
Db 25 AQAAPV 31

RESULT 22

US-09-252-991A-20273
; Sequence 20273, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20273
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20273

Query Match 1.7%; Score 7; DB 4; Length 155;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 ASTATTA 90
|||||
Db 3 ASTATTA 9

RESULT 23

US-09-252-991A-18088
; Sequence 18088, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18088
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18088

Query Match 1.7%; Score 7; DB 4; Length 159;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 120 TAETAPP 126
|||||
Db 89 TAETAPP 95

RESULT 24

US-09-252-991A-26223
; Sequence 26223, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26223
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26223

Query Match 1.7%; Score 7; DB 4; Length 179;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 134 PAATVAA 140
|||||
Db 95 PAATVAA 101

RESULT 25

US-08-974-022-52
; Sequence 52, Application US/08974022
; Patent No. 6015938

; GENERAL INFORMATION:

; APPLICANT: Boyle, William J.
; APPLICANT: Lacey, David L.
; APPLICANT: Calzone, Frank J.
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: OSTEOPROTEGERIN
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Denavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,022
; FILING DATE: 12-DEC-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/577,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-378
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-974-022-52

Query Match 1.7%; Score 7; DB 3; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 332 PEGGPG 338

```

Db      175 PEGGPGG 181
|||||
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,447A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-378D2
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-795-447A-52

Query Match          1.7% Score 7; DB 3; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      332 PEGGPGG 338
|||||
DB      175 PEGGPGG 181

RESULT 28
US-08-974-186-52
Sequence 52, Application US/08974186
Patent No. 6284740
GENERAL INFORMATION:
APPLICANT: Boyle, William J.
APPLICANT: Lacey, David L.
APPLICANT: Calzone, Frank J.
TITLE OF INVENTION: OSTEOPROTEGERIN
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,186
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/577,788
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-378
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-974-186-52

Query Match          1.7% Score 7; DB 3; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      332 PEGGPGG 338
|||||
DB      175 PEGGPGG 181

RESULT 27
US-08-795-447A-52
Sequence 52, Application US/08795447A
Patent No. 6284728
GENERAL INFORMATION:
APPLICANT: Boyle, William J.
APPLICANT: Lacey, David L.
APPLICANT: Calzone, Frank J.
APPLICANT: Chang, Ming-Shi
TITLE OF INVENTION: Osteoprotegerin
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: One Amgen Center Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91362-1789

```

Best Local Similarity 100.0%; Pred. No. 1.2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;

QY 332 PEGGPGG 338
|||||
Db 175 PEGGPGG 181

RESULT 29

US-08-795-446B-52
; Sequence 52, Application US/08795446B
; Patent No. 6288032
; GENERAL INFORMATION:
; APPLICANT: Boyle, William J.
; APPLICANT: Lacey, David L.
; APPLICANT: Calzone, Frank J.
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: OSTEOPROTEGERIN
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,446B
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/577,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-378
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-795-446B-52

Query Match 1.7%; Score 7; DB 3; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 PEGGPGG 338
|||||
Db 175 PEGGPGG 181

RESULT 30

US-08-706-945D-139
; Sequence 139, Application US/08706945D
; Patent No. 6369027
; GENERAL INFORMATION:
; APPLICANT: Boyle, William
; APPLICANT: Lacey, David
; APPLICANT: Calzone, Frank
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: Osteoprotegerin
; FILE REFERENCE: A-378CIP
; CURRENT APPLICATION NUMBER: US/08/706,945D
; CURRENT FILING DATE: 1996-09-03
; PRIOR APPLICATION NUMBER: 08/577,788

; PRIOR FILING DATE: 1995-12-22
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 139
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-706-945D-139

Query Match 1.7%; Score 7; DB 4; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 PEGGPGG 338
|||||
Db 175 PEGGPGG 181

RESULT 31

US-09-904-615-155
; Sequence 155, Application US/09904615
; Patent No. 6566325
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 155
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-904-615-155

Query Match 1.7%; Score 7; DB 4; Length 221;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 269 PGAGSGA 275
|||||
Db 35 PGAGSGA 41

RESULT 32

US-08-236-918A-6
; Sequence 6, Application US/08236918A
; Patent No. 5674704
; GENERAL INFORMATION:
; APPLICANT: Alderson, Mark R.
; APPLICANT: Goodwin, Raymond G.
; APPLICANT: Smith, Craig A.
; TITLE OF INVENTION: Cytokine Designated 4-1BB Ligand
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kathryn A. Anderson, Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple 7.5.3
; SOFTWARE: Microsoft Word, Version #6.0.1

/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/236,918A
/ FILING DATE: 06-May-1994
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/060,843
/ FILING DATE: 07-May-1993
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Anderson, Kathryn A.
/ REGISTRATION NUMBER: 32,172
/ REFERENCE/DOCKET NUMBER: 2801-B
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (206) 587-0430
/ TELEFAX: (206) 233-0644
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 256 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-236-918A-6

Query Match 1.7%; Score 7; DB 1; Length 256;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 PEGGPGG 338
DB 175 PEGGPGG 181

RESULT 33
US-09-150-864A-6
/ Sequence 6, Application US/09150864A
/ Patent No. 6355779
/ GENERAL INFORMATION:
/ APPLICANT: Alderson, Mark R.
/ APPLICANT: Goodwin, Raymond G.
/ APPLICANT: Smith, Craig A.
/ TITLE OF INVENTION: Cytokine Designated 4-1BB Ligand and Human Receptor
/ FILE REFERENCE: 2801-B
/ CURRENT APPLICATION NUMBER: US/09/150,864A
/ CURRENT FILING DATE: 1998-09-10
/ PRIOR APPLICATION NUMBER: 08/060,843
/ PRIOR FILING DATE: 1993-05-07
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 6
/ LENGTH: 256
/ TYPE: PRT
/ ORGANISM: Mus sp. (clone: mu4-1BB)
US-09-150-864A-6

Query Match 1.7%; Score 7; DB 4; Length 256;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 PEGGPGG 338
DB 175 PEGGPGG 181

RESULT 34
US-08-012-269A-2
/ Sequence 2, Application US/08012269A
/ Patent No. 6362325
/ GENERAL INFORMATION:
/ APPLICANT: Kwon, Byoung S.
/ TITLE OF INVENTION: MURINE 4-1BB GENE
/ FILE REFERENCE: 740.009U51
/ CURRENT APPLICATION NUMBER: US/08/012,269A

/ CURRENT FILING DATE: 1993-02-01
/ PRIOR APPLICATION NUMBER: US 07/922,996
/ PRIOR FILING DATE: 1992-07-30
/ PRIOR APPLICATION NUMBER: US 07/267,572
/ PRIOR FILING DATE: 1988-11-07
/ NUMBER OF SEQ ID NOS: 13
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 256
/ TYPE: PRT
/ ORGANISM: Mus musculus
US-08-012-269A-2

Query Match 1.7%; Score 7; DB 4; Length 256;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 PEGGPGG 338
DB 175 PEGGPGG 181

RESULT 35
PCT-US96-03965-2
/ Sequence 2, Application PC/TUS9603965
/ GENERAL INFORMATION:
/ APPLICANT: Kwon, Byoung Se
/ APPLICANT: Kang, Chang-Yuil
/ TITLE OF INVENTION: Monoclonal antibody against human
/ TITLE OF INVENTION: receptor 4-1BB
/ NUMBER OF SEQUENCES: 10
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Barnard, Brown & Michaels
/ STREET: 306 East State Street, Suite 220
/ CITY: Ithaca
/ STATE: NY
/ COUNTRY: USA
/ ZIP: 14850
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US96/03965
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/122,796
/ FILING DATE: 16-SEP-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/012,269
/ FILING DATE: 01-FEB-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/922,996
/ FILING DATE: 30-JUL-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/267,577
/ FILING DATE: 07-NOV-1988
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Michaels, Christopher A
/ REGISTRATION NUMBER: 34,390
/ REFERENCE/DOCKET NUMBER: KW05
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 607-273-1711
/ TELEFAX: 607-273-2609
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 256 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
PCT-US96-03965-2

Query Match 1.7%; Score 7; DB 5; Length 256;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 332 PEGPGG 338
 Db 175 PEGPGG 181
 |||||

RESULT 36

US-09-252-991A-27264
 ; Sequence 27264, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 27264
 ; LENGTH: 258
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-27264

Query Match 1.7%; Score 7; DB 4; Length 258;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 133 APAATVA 139
 Db 191 APAATVA 197
 |||||

RESULT 37

US-09-252-991A-17461
 ; Sequence 17461, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 17461
 ; LENGTH: 272
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-17461

Query Match 1.7%; Score 7; DB 4; Length 272;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 226 ASPAQPA 232
 Db 21 ASPAQPA 27
 |||||

RESULT 38

Query Match 1.7%; Score 7; DB 4; Length 281;

US-08-397-633A-75
 ; Sequence 75, Application US/08397633A
 ; Patent No. 5773577
 ; GENERAL INFORMATION:
 ; APPLICANT: Cappello, Joseph
 ; TITLE OF INVENTION: PRODUCTS COMPRISING SUBSTRATESCAPABLE
 ; TITLE OF INVENTION: OF ENZYMIC CROSS-LINKING
 ; NUMBER OF SEQUENCES: 105
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
 ; STREET: 4 Embarcadero Center, Suite 3400
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-4187
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/397,633A
 ; FILING DATE:
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Rowland, Bertram I
 ; REGISTRATION NUMBER: 20,015
 ; REFERENCE/DOCKET NUMBER: A-58848-1/BIR PROP-011-1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 781-1989
 ; TELEFAX: (415) 398-3249
 ; TELEX: 910 277299
 ; INFORMATION FOR SEQ ID NO: 75:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 281 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-397-633A-75

Query Match 1.7%; Score 7; DB 1; Length 281;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 269 PGAGSGA 275
 Db 46 PGAGSGA 52
 |||||

RESULT 39

US-09-252-991A-26761
 ; Sequence 26761, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 26761
 ; LENGTH: 281
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-26761

Best Local Similarity 100.0%; Pred. No. 1.8e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

QY 102 PVAPAAS 108
Db 61 PVAPAAS 67

RESULT 40

US-08-397-633A-76
; Sequence 76, Application US/08397633A
; Patent No. 5773577
; GENERAL INFORMATION:
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: PRODUCTS COMPRISING SUBSTRATESCAPABLE
; TITLE OF INVENTION: OF ENZYMTIC CROSS-LINKING
; NUMBER OF SEQUENCES: 105
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBRACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/397,633A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A-58848-1/BIR PROP-011-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 76:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 287 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-397-633A-76

Query Match 1.7%; Score 7; DB 1; Length 287;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 269 PGAGSGA 275
Db 49 PGAGSGA 55

RESULT 41

US-09-252-991A-22293
; Sequence 22293, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22293
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22293

Query Match 1.7%; Score 7; DB 4; Length 296;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 PATLAQP 99
Db 265 PATLAQP 271

RESULT 42

US-09-495-406-17
; Sequence 17, Application US/09495406
; Patent No. 6503744
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000110US
; CURRENT APPLICATION NUMBER: US/09/495,406
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
US-09-495-406-17

Query Match 1.7%; Score 7; DB 4; Length 303;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 ILDMGGG 173
Db 106 ILDMGGG 112

RESULT 43

US-09-134-001C-3568
; Sequence 3568, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3568
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3568

Query Match 1.7%; Score 7; DB 4; Length 341;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;

```
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 KLVNKTLL 8
    |||||
Db 211 KLVNKTLL 217

RESULT 44
US-09-252-991A-24458
; Sequence 24458, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24458
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24458

Query Match 1.7%; Score 7; DB 4; Length 343;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 344 LAAAVPQ 350
    |||||
Db 228 LAAAVPQ 234

RESULT 45
US-09-107-532A-6666
; Sequence 6666, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A. Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
```

```
TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6666:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...360
; SEQUENCE DESCRIPTION: SEQ ID NO: 6666:
US-09-107-532A-6666

Query Match 1.7%; Score 7; DB 4; Length 360;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 134 PAATVAA 140
    |||||
Db 309 PAATVAA 315

RESULT 46
US-09-252-991A-26295
; Sequence 26295, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26295
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (244)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-26295

Query Match 1.7%; Score 7; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 104 APAASVA 110
    |||||
Db 291 APAASVA 297

RESULT 47
US-09-252-991A-23727
; Sequence 23727, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
```


; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 23727
 ; LENGTH: 382
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-23727

Query Match 1.7%; Score 7; DB 4; Length 382;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 208 AQPVARA 214
 |||||
 Db 109 AQPVARA 115

RESULT 48

US-09-377-557-18
 ; Sequence 18, Application US/09377557
 ; Patent No. 6297055
 ; GENERAL INFORMATION:
 ; APPLICANT: Falco, S. Carl
 ; APPLICANT: Famodu, Layo O.
 ; APPLICANT: Orozco, Emil M. Jr.
 ; TITLE OF INVENTION: Amino Acid Decarboxylases
 ; FILE REFERENCE: BB-1237
 ; CURRENT APPLICATION NUMBER: US/09/377,557
 ; CURRENT FILING DATE: 1999-08-19
 ; EARLIER APPLICATION NUMBER: 60/099,493
 ; EARLIER FILING DATE: September 8, 1998
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 18
 ; LENGTH: 391
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 US-09-377-557-18

Query Match 1.7%; Score 7; DB 3; Length 391;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 QAAPAT 137
 |||||
 Db 25 QAAPAT 31

RESULT 49

US-09-252-991A-21075
 ; Sequence 21075, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 21075
 ; LENGTH: 406
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-21075

Query Match 1.7%; Score 7; DB 4; Length 406;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 138 VAATDDA 144
 |||||
 Db 185 VAATDDA 191

RESULT 50

US-09-252-991A-27669
 ; Sequence 27669, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 27669
 ; LENGTH: 415
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-27669

Query Match 1.7%; Score 7; DB 4; Length 415;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 269 PGAGSGA 275
 |||||
 Db 120 PGAGSGA 126

RESULT 51

US-09-252-991A-22253
 ; Sequence 22253, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 22253
 ; LENGTH: 420
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-22253

Query Match 1.7%; Score 7; DB 4; Length 420;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 342 GOLAAAV 348
 |||||
 Db 213 GOLAAAV 219

RESULT 52

US-09-252-991A-23618
 ; Sequence 23618, Application US/09252991A
 ; Patent No. 6551795

GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT FILING DATE: 1998-02-18
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23618
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23618

Query Match 1.7%; Score 7; DB 4; Length 423;

Best Local Similarity 100.0%; Pred. No. 2.6e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 343 QLAAPV 349
Db 212 QLAAPV 218

RESULT 53
US-08-476-509B-2
; Sequence 2, Application US/08476509B
; Patent No. 6034212
; GENERAL INFORMATION:
; APPLICANT: SUDOL, MARIUS
; APPLICANT: PEER, BORK
; APPLICANT: HENRY, CHEN
; TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A
; TITLE OF INVENTION: SIGNALING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE
; TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,509B
; FILING DATE: 01-DEC-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-101 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1694
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 448 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-476-509B-2

Query Match 1.7%; Score 7; DB 3; Length 448;

Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 269 PGAGSGA 275
Db 28 PGAGSGA 34
RESULT 54
US-08-948-997-6
; Sequence 6, Application US/08948997
; Patent No. 6008020
; GENERAL INFORMATION:
; APPLICANT: HASTINGS, GREGG
; APPLICANT: COLEMAN, TIM
; TITLE OF INVENTION: BRAIN-ASSOCIATED INHIBITOR OF
; TITLE OF INVENTION: TISSUE-TYPE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,997
; FILING DATE: Oct-10-97
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: A. ANDERS BROOKES
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF336
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 465 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-948-997-6

Query Match 1.7%; Score 7; DB 3; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 269 PGAGSGA 275
Db 4 PGAGSGA 10

RESULT 55
US-09-348-817A-6
; Sequence 6, Application US/09348817A
; Patent No. 6191260
; GENERAL INFORMATION:
; APPLICANT: Hastings et al.
; TITLE OF INVENTION: Brain-Associated Inhibitor of Tissue-Type Plasminogen
; TITLE OF INVENTION: Activator
; FILE REFERENCE: PF336D1
; CURRENT APPLICATION NUMBER: US/09/348,817A
; CURRENT FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 08/948,997
; PRIOR FILING DATE: 1997-10-10

; PRIOR APPLICATION NUMBER: 60/028,117
; PRIOR FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-348-817A-6

Query Match 1.7%; Score 7; DB 3; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 269 PGAGSGA 275
Db 4 PGAGSGA 10

RESULT 56
US-09-722-292-6
; Sequence 6, Application US/09722292
; Patent No. 6541452
; GENERAL INFORMATION:
; APPLICANT: Hastings et al.
; TITLE OF INVENTION: Brain-Associated Inhibitor of Tissue-Type Plasminogen
; TITLE OF INVENTION: Activator
; FILE REFERENCE: PF336D1
; CURRENT APPLICATION NUMBER: US/09/722,292
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/348,817
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/028,117
; PRIOR FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-722-292-6

Query Match 1.7%; Score 7; DB 4; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 269 PGAGSGA 275
Db 4 PGAGSGA 10

RESULT 57
US-09-252-991A-23144
; Sequence 23144, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23144
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23144

Query Match 1.7%; Score 7; DB 4; Length 479;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 QPAAAPVA 104
Db 302 QPAAAPVA 308

RESULT 58
US-08-348-518C-2
; Sequence 2, Application US/08348518C
; Patent No. 6022740
; GENERAL INFORMATION:
; APPLICANT: SUDOL, MARIUS
; APPLICANT: PEER, BORK
; APPLICANT: HENRY, CHEN
; TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A
; TITLE OF INVENTION: SIGNALLING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE
; TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC USES
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/348,518C
; FILING DATE: 01-DEC-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 486 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-348-518C-2

Query Match 1.7%; Score 7; DB 3; Length 486;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 269 PGAGSGA 275
Db 28 PGAGSGA 34

RESULT 59
US-09-252-991A-27501
; Sequence 27501, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27501
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27501

Query Match 1.7%; Score 7; DB 4; Length 495;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 AKASSG 83
Db 458 AKASSG 464

RESULT 60

US-09-252-991A-16737
; Sequence 16737, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16737
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16737

Query Match 1.7%; Score 7; DB 4; Length 519;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 213 RAPAAQG 219
Db 2 RAPAAQG 8

RESULT 61

US-09-252-991A-22537
; Sequence 22537, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22537
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22537

Query Match 1.7%; Score 7; DB 4; Length 534;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 AAPVAPA 106
Db 267 AAPVAPA 273

RESULT 62

US-09-252-991A-26327
; Sequence 26327, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26327
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26327

Query Match 1.7%; Score 7; DB 4; Length 549;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 AASNLVF 156
Db 40 AASNLVF 46

RESULT 63

US-09-252-991A-23080
; Sequence 23080, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23080
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23080

Query Match 1.7%; Score 7; DB 4; Length 561;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 PVQPSA 242
Db 39 PVQPSA 45

RESULT 64

US-09-205-258-573
; Sequence 573, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 573
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: SITE
; LOCATION: (409)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-205-258-573

Query Match 1.7%; Score 7; DB 4; Length 567;
Best Local Similarity 100.0%; Pred.No. 3.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 GASTATT 89
Db 79 GASTATT 85

RESULT 65
US-09-199-290-7
; Sequence 7, Application US/09199290
; Patent No. 6235084
; GENERAL INFORMATION:
; APPLICANT: Nielsen, Bjarne R.
; APPLICANT: Nielsen, Ruby
; APPLICANT: Lehmebeck, Jan
; TITLE OF INVENTION: Thermostable Glucoamylase
; FILE REFERENCE: 5279.200-US
; CURRENT APPLICATION NUMBER: US/09/199,290
; CURRENT FILING DATE: 1998-11-24
; EARLIER APPLICATION NUMBER: 1557/97
; EARLIER FILING DATE: 1997-12-30
; EARLIER APPLICATION NUMBER: 0925/98
; EARLIER FILING DATE: 1998-07-10
; EARLIER APPLICATION NUMBER: 60/070,746
; EARLIER FILING DATE: 1998-01-08
; EARLIER APPLICATION NUMBER: 60/094,344
; EARLIER FILING DATE: 1998-07-28
; EARLIER APPLICATION NUMBER: 08/979,673
; EARLIER FILING DATE: 1997-11-26
; EARLIER APPLICATION NUMBER: 09/107,657
; EARLIER FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Talaromyces emersonii
US-09-199-290-7

Query Match 1.7%; Score 7; DB 3; Length 591;
Best Local Similarity 100.0%; Pred.No. 3.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 160 LEQTIQ 166
|||||
Db 76 LEQTIQ 82

RESULT 66
US-09-252-991A-23472
; Sequence 23472, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23472
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23472

Query Match 1.7%; Score 7; DB 4; Length 617;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 275 ALDALRQ 281
|||||
Db 341 ALDALRQ 347

RESULT 67
US-09-199-290-34
; Sequence 34, Application US/09199290
; Patent No. 6255084
; GENERAL INFORMATION:
; APPLICANT: Nielsen, Bjarne R.
; APPLICANT: Nielsen, Ruby
; APPLICANT: Lehmebeck, Jan
; TITLE OF INVENTION: Thermostable Glucoamylase
; FILE REFERENCE: 5279.200-US
; CURRENT APPLICATION NUMBER: US/09/199,290
; CURRENT FILING DATE: 1998-11-24
; EARLIER APPLICATION NUMBER: 1557/97
; EARLIER FILING DATE: 1997-12-30
; EARLIER APPLICATION NUMBER: 0925/98
; EARLIER FILING DATE: 1998-07-10
; EARLIER APPLICATION NUMBER: 60/070,746
; EARLIER FILING DATE: 1998-01-08
; EARLIER APPLICATION NUMBER: 60/094,344
; EARLIER FILING DATE: 1998-07-28
; EARLIER APPLICATION NUMBER: 08/979,673
; EARLIER FILING DATE: 1997-11-26
; EARLIER APPLICATION NUMBER: 09/107,657
; EARLIER FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Talaromyces emersonii
US-09-199-290-34

Query Match 1.7%; Score 7; DB 3; Length 618;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 160 LEQTIQ 166

Db 103 LEQTIQ 109
|||||

RESULT 68
US-09-252-991A-21585
; Sequence 21585, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21585
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21585

Query Match 1.7%; Score 7; DB 4; Length 619;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 138 VAATDDA 144
|||||
Db 589 VAATDDA 595

RESULT 69
US-09-252-991A-28001
; Sequence 28001, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28001
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28001

Query Match 1.7%; Score 7; DB 4; Length 657;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 97 AQPAPV 103
|||||
Db 538 AQPAPV 544

RESULT 70
US-08-795-868-14
; Sequence 14, Application US/08795868
; Patent No. 5846773
; GENERAL INFORMATION:
; APPLICANT: Lee, Mu-En
; APPLICANT: Hsieh, Chung-Ming

/ TITLE OF INVENTION: A SINGLE GENE ENCODING AORTIC-SPECIFIC
/ TITLE OF INVENTION: AND STRIATED-SPECIFIC MUSCLE CELL ISOFORMS AND USES THEREOF
/ NUMBER OF SEQUENCES: 19
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Fish & Richardson, P.C.
/ STREET: 225 Franklin Street
/ CITY: Boston
/ STATE: MA
/ COUNTRY: US
/ ZIP: 02110-2804
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/795,868
/ FILING DATE: 06-FEB-1997
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/494,577
/ FILING DATE: 22-JUN-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fraser, Janis K.
/ REGISTRATION NUMBER: 34,819
/ REFERENCE/DOCKET NUMBER: 05433/032001
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-542-5070
/ TELEFAX: 617-542-8906
/ INFORMATION FOR SEQ ID NO: 14:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 661 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: not relevant
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ FRAGMENT TYPE: Internal
/ US-08-795-868-14

Query Match 1.7%; Score 7; DB 2; Length 661;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 346 AAVPQTL 352
Db 350 AAVPQTL 356
|||||

RESULT 71
US-09-303-069-14
/ Sequence 14, Application US/09303069A
/ Patent No. 6350592
/ GENERAL INFORMATION:
/ APPLICANT: Lee, Mu-En
/ APPLICANT: Hsieh, Chung-Ming
/ TITLE OF INVENTION: SINGLE GENE ENCODING AORTIC-SPECIFIC AND STRIATED-SPECIFIC
/ TITLE OF INVENTION: MUSCLE CELL ISOFORMS AND USES THEREOF
/ FILE REFERENCE: 05433/039001
/ CURRENT APPLICATION NUMBER: US/09/303,069A
/ CURRENT FILING DATE: 1999-04-30
/ EARLIER APPLICATION NUMBER: US 09/134,250
/ EARLIER FILING DATE: 1998-08-14
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 14
/ LENGTH: 661
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-09-303-069-14

Query Match 1.7%; Score 7; DB 4; Length 661;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 346 AAVPQTL 352
Db 350 AAVPQTL 356
|||||

RESULT 72
US-09-134-250-14
/ Sequence 14, Application US/09134250B
/ Patent No. 6399753
/ GENERAL INFORMATION:
/ APPLICANT: Lee, Mu-En
/ APPLICANT: Hsieh, Chung-Ming
/ TITLE OF INVENTION: SINGLE GENE ENCODING AORTIC-SPECIFIC AND STRIATED-SPECIFIC
/ TITLE OF INVENTION: MUSCLE CELL ISOFORMS AND USES THEREOF
/ FILE REFERENCE: 05433/038001
/ CURRENT APPLICATION NUMBER: US/09/134,250B
/ CURRENT FILING DATE: 1998-08-14
/ EARLIER APPLICATION NUMBER: US 08/795,868
/ EARLIER FILING DATE: 1997-02-06
/ EARLIER APPLICATION NUMBER: US 08/494,577
/ EARLIER FILING DATE: 1995-06-22
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 14
/ LENGTH: 661
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-09-134-250-14

Query Match 1.7%; Score 7; DB 4; Length 661;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 346 AAVPQTL 352
Db 350 AAVPQTL 356
|||||

RESULT 73
US-09-252-991A-20752
/ Sequence 20752, Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ CURRENT FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 20752
/ LENGTH: 749
/ TYPE: PRT
/ ORGANISM: Pseudomonas aeruginosa
/ US-09-252-991A-20752

Query Match 1.7%; Score 7; DB 4; Length 749;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 275 ALDALRQ 281
Db 446 ALDALRQ 452
|||||

RESULT 74
US-09-252-991A-23774
/ Sequence 23774, Application US/09252991A

; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23774
; LENGTH: 798
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23774

Query Match 1.7%; Score 7; DB 4; Length 798;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 104 APAASVA 110
Db 243 APAASVA 249

RESULT 75
US-09-252-991A-25050
; Sequence 25050, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25050
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25050

Query Match 1.7%; Score 7; DB 4; Length 802;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 212 ARAPAAG 218
Db 525 ARAPAAG 531

Search completed: December 17, 2003, 06:36:28
Job time : 17.1462 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 06:27:31 ; Search time 16.6624 Seconds
(without alignments)
2123.954 Million cell updates/sec

Title: US-09-805-550-4

Perfect score: 368

Sequence: 1 MKLVTKLKGTHFEIRVQPN.....CORNEELAANYLLEHAGEED 368

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : PIR_76.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	21	5.7	367	2	F96827	protein F20B17.8 [
2	15	4.1	113	2	H86296	P309.1 protein - A
3	15	4.1	379	2	T14337	RAD23 protein, iso
4	15	4.1	382	2	T14336	RAD23 protein, iso
5	14	3.8	246	2	G86296	T24D18.27 protein
6	13	3.5	392	2	T04150	RAD23 protein homo
7	9	2.4	225	2	H75381	DNA-binding respon
8	9	2.4	368	2	T40115	uv excision repair
9	9	2.4	753	2	A96747	probable RNA-bindi
10	8	2.2	111	2	T14306	glycine-rich prote
11	8	2.2	173	2	JQ1064	glycine-rich prote
12	8	2.2	180	2	T49530	related to glycine
13	8	2.2	210	2	JQ1060	glycine-rich prote
14	8	2.2	220	2	T14441	glycine-rich prote
15	8	2.2	276	2	B25345	tropoin 1, cardia
16	8	2.2	338	1	KNMU	glycine-rich cell
17	8	2.2	384	1	A26099	glycine-rich cell
18	8	2.2	401	2	C96614	hypothetical prote
19	8	2.2	465	1	S01820	glycine-rich cell
20	8	2.2	548	2	T28910	hypothetical prote
21	8	2.2	631	2	G64874	probable membrane
22	8	2.2	712	2	A45638	immunodominant mic
23	8	2.2	809	2	S55344	outer envelope mem
24	8	2.2	823	2	T02812	probable membrane
25	8	2.2	1010	2	I40329	brkA proteoin - Bor
26	8	2.2	1454	2	T13709	diacylglycerol kin
27	8	2.2	1718	2	T14603	hypothetical prote
28	8	2.2	2425	2	D69426	surface layer prot
29	8	2.2	3016	2	S77300	hypothetical prote

30	7	1.9	42	2	A71429	hypothetical prote
31	7	1.9	50	2	B49410	t-complex polypept
32	7	1.9	76	2	A39034	neurogranin - bovi
33	7	1.9	78	2	A57288	neurogranin - rat
34	7	1.9	78	2	I47043	neurogranin RC3 [i
35	7	1.9	94	2	S49470	hypothetical prote
36	7	1.9	96	2	F97355	uncharacterized co
37	7	1.9	102	2	S69752	hypothetical prote
38	7	1.9	102	2	T45500	hypothetical prote
39	7	1.9	108	2	G86252	hypothetical prote
40	7	1.9	112	2	C55993	hypothetical prote
41	7	1.9	119	2	S58445	lipoprotein D - Sa
42	7	1.9	119	2	D83723	hypothetical prote
43	7	1.9	121	2	A24272	ig heavy chain pre
44	7	1.9	172	2	T36107	probable serine/ar
45	7	1.9	180	2	C82243	conserved hypothet
46	7	1.9	182	2	I59203	gene HOX2.8 protei
47	7	1.9	195	2	AF0830	probable membrane
48	7	1.9	201	2	T00799	hypothetical prote
49	7	1.9	201	2	G01204	twist protein homo
50	7	1.9	206	2	I53066	gene M-twist prote
51	7	1.9	207	2	T07854	germin-like protei
52	7	1.9	209	2	H69037	conserved hypothet
53	7	1.9	215	2	C87542	hypothetical prote
54	7	1.9	216	2	T48482	hypothetical prote
55	7	1.9	218	2	AF3273	ribosomal protein
56	7	1.9	224	2	A53143	testis-determining
57	7	1.9	228	2	A98166	hypothetical prote
58	7	1.9	228	2	AF3121	hypothetical prote
59	7	1.9	231	2	T07358	ribosomal protein
60	7	1.9	234	2	G95989	hypothetical glyci
61	7	1.9	239	2	I46082	CD8 alpha-chain -
62	7	1.9	247	2	T46968	flavoprotein [impo
63	7	1.9	248	2	T33230	hypothetical prote
64	7	1.9	249	2	B33144	homeotic protein U
65	7	1.9	255	2	B84777	hypothetical prote
66	7	1.9	266	1	CIRBL	calpain [SC 3.4.22
67	7	1.9	268	2	S09860	hypothetical prote
68	7	1.9	272	1	G83023	3',5'-cyclic-nucle
69	7	1.9	274	2	A83583	probable biotin sy
70	7	1.9	276	2	AH2885	hypothetical prote
71	7	1.9	276	2	E97651	dipeptide transpor
72	7	1.9	284	2	S74256	homeotic protein s
73	7	1.9	290	2	T23416	hypothetical prote
74	7	1.9	291	1	S31415	glycine-rich prote
75	7	1.9	309	2	T00503	probable MYB famli
76	7	1.9	320	2	F82763	D-alanine-D-alanin
77	7	1.9	323	2	S16318	homeotic protein H
78	7	1.9	325	2	A75415	homoserine dehydro
79	7	1.9	327	2	B75341	probable oxidoredu
80	7	1.9	329	2	A37864	SCL protein homolo
81	7	1.9	330	2	S74255	homeotic protein s
82	7	1.9	333	2	AF0407	lipoprotein [impor
83	7	1.9	339	2	F97190	phenylalanyl-tRNA
84	7	1.9	341	2	AG1210	glucitol dehydroge
85	7	1.9	342	2	S14432	heterogeneous ribo
86	7	1.9	344	2	T33057	hypothetical prote
87	7	1.9	344	2	T48827	hypothetical prote
88	7	1.9	346	2	B98222	probable transcrip
89	7	1.9	347	2	AF3064	transcription regu
90	7	1.9	348	2	AH0090	probable regulator
91	7	1.9	362	2	T29278	hypothetical prote
92	7	1.9	366	2	T44827	probable outer mem
93	7	1.9	366	2	T27309	hypothetical prote
94	7	1.9	367	2	JC6087	helix-loop-helix t
95	7	1.9	373	2	AH0855	lipoprotein NlpD p
96	7	1.9	374	2	T33173	hypothetical prote
97	7	1.9	375	2	E96567	hypothetical prote
98	7	1.9	376	2	AD1621	heat shock protein
99	7	1.9	377	2	T43739	heat shock protein
100	7	1.9	377	2	AH1258	heat shock protein

ALIGNMENTS

RESULT 1

F96827
 Protein F20B17.8 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002
 C/Accession: F96827
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federici, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A/Accession: F96827
 A>Status: preliminary
 A:Molecule type: DNA
 A/Residues: 1-367 <STO>
 A/Cross-references: GB:AE005173; NID:g7715605; PIDN:AAF68123.1; GSPDB:GN00141
 C/Genetics:
 A:Gene: F20B17.8
 A:Map position: 1
 C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

Query Match 5.7%; Score 21; DB 2; Length 367;
 Best Local Similarity 100.0%; Pred. No. 5.6e-13;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 342 VIEAFLACDRNEELAAAYLLE 362
 Db 340 VIEAFLACDRNEELAAAYLLE 360

RESULT 2

H86296
 F309.1 protein - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
 C/Accession: H86296
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federici, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A/Accession: H86296
 A>Status: preliminary
 A:Molecule type: DNA
 A/Residues: 1-113 <STO>
 A/Cross-references: GB:AE005172; NID:g4966345; PIDN:AAD34676.1; GSPDB:GN00141
 C/Genetics:
 A:Map position: 1

Query Match 4.1%; Score 15; DB 2; Length 113;
 Best Local Similarity 100.0%; Pred. No. 2.4e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 349 CDNEELAAAYLLEH 363
 Db 93 CDNEELAAAYLLEH 107

RESULT 3

T14337
 RAD23 protein, isoform II - carrot
 C:Species: Daucus carota (carrot)
 C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
 C/Accession: T14337
 R:Sturm, A.; Leinhardt, S.
 Plant J. 13, 815-821, 1998
 A>Title: Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in yeast.
 A:Reference number: Z17989; MUID:98345997; PMID:9681019
 A/Accession: T14337
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A/Residues: 1-379 <STU>
 A/Cross-references: EMBL:Y12014; NID:g1914684; PIDN:CAA72742.1; PID:g1914685
 A/Experimental source: subspecies Queen Anne's lace, isolate W001C
 C/Genetics:
 A:Gene: RAD23-2
 C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

Query Match 4.1%; Score 15; DB 2; Length 379;
 Best Local Similarity 100.0%; Pred. No. 7.4e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 342 VIEAFLACDRNEELA 356

Db 352 VIEAFLACDRNEELA 366

RESULT 4

T14336
 RAD23 protein, isoform I - carrot
 C:Species: Daucus carota (carrot)
 C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
 C/Accession: T14336
 R:Sturm, A.; Leinhardt, S.
 Plant J. 13, 815-821, 1998
 A>Title: Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in yeast.
 A:Reference number: Z17989; MUID:98345997; PMID:9681019
 A/Accession: T14336
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A/Residues: 1-382 <STU>
 A/Cross-references: EMBL:Y12013; NID:g1914682; PIDN:CAA72741.1; PID:g1914683
 A/Experimental source: subspecies Queen Anne's lace, isolate W001C
 C/Genetics:
 A:Gene: RAD23-1
 C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

Query Match 4.1%; Score 15; DB 2; Length 382;
 Best Local Similarity 100.0%; Pred. No. 7.5e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 NNPERAVEYLXSGIP 189

Db 187 NNPERAVEYLXSGIP 201

RESULT 5

G86296
 T24D18.27 protein - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
 C/Accession: G86296
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federici, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: G86296
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-246 <STO>
 A:Cross-references: GB:AE005172; NID:G6587822; PIDN:AAF18513.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1

Query Match 3.8%; Score 14; DB 2; Length 246;
 Best Local Similarity 100.0%; Pred. No. 5.2e-06; Mismatches 0; Indels 0; Gaps 0;
 Matches 14; Conservative 0;

Qy 167 RALRAAYNNPERAV 180
 |||||
 Db 155 RALRAAYNNPERAV 168

RESULT 6
 T04150
 RAD23 protein homolog - rice
 C:Species: Oryza sativa (rice)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
 C:Accession: T04150
 R:Schultz, T.F.; Quatrano, R.S.
 Plant Mol. Biol. 34, 557-562, 1997
 A>Title: Characterization and expression of a rice RAD23 gene.
 A:Reference number: Z08695; MUID:97369378; PMID:9225866
 A:Accession: T04150
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-392 <SCH>
 A:Cross-references: EMBL:U63530; NID:g1488296; PIDN:AAB65841.1; PID:g1488297
 A:Experimental source: cv. Nipponbare
 C:Genetics:
 A:Gene: RAD23
 C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

Query Match 3.5%; Score 13; DB 2; Length 392;
 Best Local Similarity 100.0%; Pred. No. 8.3e-05; Mismatches 0; Indels 0; Gaps 0;
 Matches 13; Conservative 0;

Qy 174 NNPERAVEVLYSG 186
 |||||
 Db 187 NNPERAVEVLYSG 199

RESULT 7
 H75381
 DNA-binding response regulator - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 12-May-2003
 C:Accession: H75381
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: H75381
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-225 <WHI>
 A:Cross-references: GB:AE001999; GB:AE000513; NID:g6459316; PIDN:AAF11120.1; PID:g645932
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR1558
 A:Map position: 1
 C:Superfamily: response regulator with an HTH DNA-binding domain, NarL type; response re

Query Match 2.4%; Score 9; DB 2; Length 225;
 Best Local Similarity 100.0%; Pred. No. 0.59;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 57 STLEENKVN 65
 |||||
 Db 95 STLEENKVN 103

RESULT 8
 T40115
 uv excision repair protein rad23 homolog - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Sep-2000
 C:Accession: T40115; T51298
 R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Taylor, K.; Harris, D.
 submitted to the EMBL Data Library, September 1998
 A:Reference number: Z21906
 A:Accession: T40115
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-368 <WOO>
 A:Cross-references: EMBL:AL031788; PIDN:CAA21170.1; GSPDB:GN00067; SPDB:SPBC2D10.12
 A:Experimental source: strain 972h-; cosmid c2D10
 R:Zhao, Y.; Elder, R.T.
 submitted to the EMBL Data Library, July 1999
 A:Description: A fission yeast orthologue (rhp23) of the human nucleotide excision rep
 A:Reference number: Z25362
 A:Accession: T51298
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-368 <ZHA>
 A:Cross-references: EMBL:AF174293; PIDN:AAD51975.1
 A:Experimental source: strain SP223
 C:Genetics:
 A:Gene: rhp23; SPDB:SPBC2D10.12
 A:Map position: 2
 A:Introns: 23/3; 48/1; 328/3; 351/3
 C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

Query Match 2.4%; Score 9; DB 2; Length 368;
 Best Local Similarity 100.0%; Pred. No. 0.93; Mismatches 0; Indels 0; Gaps 0;
 Matches 9; Conservative 0;

Qy 352 NEELAANYL 360
 |||||
 Db 349 NEELAANYL 357

RESULT 9
 A96747
 Probable RNA-binding domain T10D10.21 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: A96747
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alons
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 815-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzial
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: A96747
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-753 <STO>
 A:Cross-references: GB:AE005173; NID:g6730769; PIDN:AAF27158.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: T10D10.21
 A:Map position: 1

Query Match 2.4%; Score 9; DB 2; Length 753;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 340 ARVIEAFLA 348
 |||||
 Db 542 ARVIEAFLA 550

RESULT 10

Tl4306
 glycine-rich protein - carrot (fragment)
 C:Species: Daucus carota (carrot)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: Tl4306
 R:Lin, X.; Hwang, G.J.; Zimmerman, J.L.
 submitted to the EMBL Data Library, January 1996
 A:Description: Isolation and characterization of a diverse set of genes from carrot somatic cells
 A:Reference number: Z17968
 A:Accession: Tl4306
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-111 <LIN>
 A:Cross-references: EMBL:U47097; NID:g1276970; PID:g1276971
 A:Experimental source: strain Danver Half-long

Query Match 2.2%; Score 8; DB 2; Length 111;
 Best Local Similarity 100.0%; Pred. No. 3.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGGG 241
 |||||
 Db 18 AGGAGGG 25

RESULT 11

QJ01064
 glycine-rich protein 5 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 20-Aug-1999
 C:Accession: JQ1064
 R:de Oliveira, D.E.; Seurinck, J.; Inze, D.; Van Montagu, M.; Botterman, J.
 Plant Cell 2, 427-436, 1990
 A:Title: Differential expression of five Arabidopsis genes encoding glycine-rich protein
 A:Reference number: JQ1060; MUID:93044485; PMID:2152168
 A:Molecule type: mRNA
 A:Residues: 1-173 <DEO>
 A:Cross-references: GB:S47414; NID:g259450; PIDN:AAB24077.1; PID:g259451
 A:Experimental source: strain C24
 C:Superfamily: glycine-rich cell wall structural protein 1

Query Match 2.2%; Score 8; DB 2; Length 173;
 Best Local Similarity 100.0%; Pred. No. 4.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGGG 241
 |||||
 Db 84 AGGAGGG 91

RESULT 12

T49530
 related to glycine-rich cell wall structural protein [imported] - Neurospora crassa
 N:Alternate names: protein B21J21.90
 C:Species: Neurospora crassa
 C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
 C:Accession: T49530
 R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, R.
 submitted to the Protein Sequence Database, May 2000
 A:Reference number: Z25022
 A:Accession: T49530
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-180 <SCH>
 A:Cross-references: EMBL:AL355929; GSPDB:GN00116; NCSP:B21J21.90
 A:Experimental source: BAC clone B21J21; strain OR74A
 C:Genetics:
 A:Gene: NCSP:B21J21.90
 A:Map position: 6

Query Match 2.2%; Score 8; DB 2; Length 180;
 Best Local Similarity 100.0%; Pred. No. 5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGGG 241
 |||||
 Db 60 AGGAGGG 67

RESULT 13

JQ1060
 glycine-rich protein 1 - Arabidopsis thaliana (fragment)
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 20-Aug-1999
 C:Accession: JQ1060
 R:de Oliveira, D.E.; Seurinck, J.; Inze, D.; Van Montagu, M.; Botterman, J.
 Plant Cell 2, 427-436, 1990
 A:Title: Differential expression of five Arabidopsis genes encoding glycine-rich protein
 A:Reference number: JQ1060; MUID:93044485; PMID:2152168
 A:Molecule type: mRNA
 A:Residues: 1-210 <DEO>
 A:Cross-references: GB:S47405; NID:g259442; PIDN:AAB24073.1; PID:g259443
 A:Experimental source: strain C24
 C:Superfamily: Phaseolus glycine-rich cell wall protein 1.8

Query Match 2.2%; Score 8; DB 2; Length 210;
 Best Local Similarity 100.0%; Pred. No. 5.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGGG 241
 |||||
 Db 148 AGGAGGG 155

RESULT 14

Tl4441
 glycine-rich protein - wild cabbage (fragment)
 C:Species: Brassica oleracea (wild cabbage)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Jun-2000
 C:Accession: Tl4441
 R:Booker, J.P.; Crov, R.D.
 submitted to the EMBL Data Library, July 1996
 A:Reference number: Z18090
 A:Accession: Tl4441
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
 A:Residues: 1-220 <BOO>
 A:Cross-references: EMBL:Z74892
 A:Experimental source: stigma
 C:Superfamily: Phaseolus glycine-rich cell wall protein 1.8

Query Match 2.2%; Score 8; DB 2; Length 220;
 Best Local Similarity 100.0%; Pred. No. 6.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGGG 241
 |||||
 Db 16 AGGAGGG 23

RESULT 15

B25345
 tropinin T, cardiac muscle, minor isoform - rabbit
 C:Species: Oryctolagus cuniculus (domestic rabbit)

C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 23-Feb-1997

C;Accession: B25345

R;Pearlatone, J.R.; Carpenter, M.R.; Smillie, L.B.

J. Biol. Chem. 261, 16795-16810, 1986

A;Title: Amino acid sequence of rabbit cardiac troponin T.

A;Reference number: A92569; MUID:87057385; PMID:3782144

A;Accession: B25345

A;Molecule type: protein

A;Residues: 1-276 <PEA>

C;Superfamily: troponin T

C;Keywords: alternative splicing; cardiac muscle; differentiation; heart; muscle; phosph

Query Match 2.2%; Score 8; DB 2; Length 276;

Best Local Similarity 100.0%; Pred. No. 7.5;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGGG 241

Db 21 AGGGAGGG 28

RESULT 16

KNMU

glycine-rich cell wall protein precursor - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jul-1999

C;Accession: S17732

R;Quigley, F.; Villiot, M.L.; Mache, R.

Plant Mol. Biol. 17, 949-952, 1991

A;Title: Nucleotide sequence and expression of a novel glycine-rich protein gene from A

A;Reference number: S17732; MUID:92003708; PMID:1912511

A;Accession: S17732

A;Molecule type: DNA

A;Residues: 1-338 <QUI>

A;Cross-references: EMBL:X58338; NID:gl6292; PIDN:CAA41249.1; PID:gl6293

C;Genetics:

A;Introns: 203/1

C;Superfamily: Phaseolus glycine-rich cell wall protein 1.8

C;Keywords: cell wall; structural protein

F;1-20/Domain: signal sequence #status predicted <SIG>

F;21-338/Product: glycine-rich cell wall protein #status predicted <MAT>

Query Match 2.2%; Score 8; DB 1; Length 338;

Best Local Similarity 100.0%; Pred. No. 9;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGGG 241

Db 52 AGGGAGGG 59

RESULT 17

A26099

glycine-rich cell wall structural protein - garden petunia

C;Species: Petunia x hybrida (garden petunia)

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C;Accession: A26099

R;Condit, C.M.; Meagher, R.B.

Nature 323, 178-181, 1986

A;Title: A gene encoding a novel glycine-rich structural protein of petunia.

A;Reference number: A26099

A;Accession: A26099

A;Molecule type: mRNA

A;Residues: 1-384 <CON>

C;Superfamily: Phaseolus glycine-rich cell wall protein 1.8

Query Match 2.2%; Score 8; DB 1; Length 384;

Best Local Similarity 100.0%; Pred. No. 10;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGGG 241

Db 75 AGGGAGGG 82

RESULT 18

C96614

hypothetical protein T18124.2 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C;Accession: C96614

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alons

anzen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzial

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: C96614

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-401 <STO>

A;Cross-references: GB:AE005173; NID:gl1038497; PIDN:AG27774.1; GSPDB:GN00141

C;Genetics:

A;Gene: T18124.2

A;Map position: 1

Query Match 2.2%; Score 8; DB 2; Length 401;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGGG 241

Db 294 AGGGAGGG 301

RESULT 19

S01820

glycine-rich cell wall protein 1.8 precursor - kidney bean

C;Species: Phaseolus vulgaris (kidney bean)

C;Date: 30-Sep-1989 #sequence_revision 19-May-1994 #text_change 16-Jul-1999

C;Accession: S01820

R;Keller, B.; Sauer, N.; Lamb, C.J.

EMBO J. 7, 3625-3633, 1988

A;Title: Glycine-rich cell wall proteins in bean: gene structure and association of th

A;Reference number: S01820; MUID:89091109; PMID:3208742

A;Accession: S01820

A;Molecule type: DNA

A;Residues: 1-465 <KEL>

A;Cross-references: EMBL:X13596; NID:g21002; PIDN:CAA31932.1; PID:g21003

C;Comment: This protein is enriched in the cell wall fraction of young hypocotyls and

tyls.

C;Comment: Much of the sequence consists of tandemly repeated 22-residue segments with

C;Superfamily: Phaseolus glycine-rich cell wall protein 1.8

C;Keywords: cell wall; structural protein; tandem repeat

F;1-30/Domain: signal sequence #status predicted <SIG>

F;31-465/Product: glycine-rich cell wall protein 1.8 #status predicted <MAT>

Query Match 2.2%; Score 8; DB 1; Length 465;

Best Local Similarity 100.0%; Pred. No. 12;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGGG 241

Db 191 AGGGAGGG 198

RESULT 20

T28910

hypothetical protein T26C11.4 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000

C;Accession: T28910
 R;Martin, J.
 submitted to the EMBL Data Library, December 1995
 A;Description: The sequence of *C. elegans* cosmid T26C11.
 A;Reference number: Z20542
 A;Accession: T28910
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-548 <WAR>
 A;Cross-references: EMBL:U41017; PIDN:AAC48213.1; GSPDB:GN00028; CBSP:T26C11.4
 A;Experimental source: strain Bristol N2; clone T26C11
 C;Genetics:
 A;Gene: CBSP:T26C11.4
 A;Map position: X
 A;Introns: 9/2; 85/1; 174/1; 202/1; 229/1; 261/1; 298/1; 337/1; 370/1; 393/1; 434/1; 471/1

Query Match 2.2%; Score 8; DB 2; Length 548;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 POGASNAG 235
 |||||
 Db 195 POGASNAG 202

RESULT 21
 G64874
 probable membrane protein yciQ - *Escherichia coli* (strain K-12)
 C;Species: *Escherichia coli*
 C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
 C;Accession: G64874
 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A;Title: The complete genome sequence of *Escherichia coli* K-12.
 A;Reference number: A64720; MUID:97426617; PMID:9278503
 A;Accession: G64874
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-631 <BLAT>
 A;Cross-references: GB:AE000224; GB:U00096; NID:g1787509; PIDN:AAC74350.1; PID:g1787522;
 A;Experimental source: strain K-12, substrain MGL555
 C;Genetics:
 A;Gene: yciQ
 C;Superfamily: *Escherichia coli* probable membrane protein yciQ
 C;Keywords: transmembrane protein
 F;7-23/Domain: transmembrane #status predicted <TM1>
 F;249-265/Domain: transmembrane #status predicted <TM2>
 F;402-418/Domain: transmembrane #status predicted <TM3>
 F;431-447/Domain: transmembrane #status predicted <TM4>
 F;460-476/Domain: transmembrane #status predicted <TM5>

Query Match 2.2%; Score 8; DB 2; Length 631;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
 |||||
 Db 619 AGGGAGGG 626

RESULT 22
 A45638
 immunodominant microneme protein Etp100 - *Eimeria tenella*
 C;Species: *Eimeria tenella*
 C;Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
 C;Accession: A45638
 R;Tomley, F.M.; Clarke, L.E.; Kawazo, U.; Dijkema, R.; Kok, J.J.
 Mol. Biochem. Parasitol. 49, 277-288, 1991
 A;Title: Sequence of the gene encoding an immunodominant microneme protein of *Eimeria tenella*
 A;Reference number: A45638; MUID:92131064; PMID:1775171
 A;Accession: A45638
 A;Status: preliminary

A;Molecule type: DNA
 A;Residues: 1-712 <TOM>
 A;Cross-references: GB:AF032905; GB:M73495; NID:g2707732; PIDN:AAD03350.1; PID:g2707732;
 A;Note: sequence extracted from NCBI backbone (NCBIN:77752; NCBI:77756)
 C;Superfamily: thrombospondin type 1 repeat homology; von Willebrand factor type A repeat homology
 F;18-218/Domain: von Willebrand factor type A repeat homology <VWA1>
 F;238-296/Domain: thrombospondin type 1 repeat homology <THR1>
 F;309-371/Domain: thrombospondin type 1 repeat homology <THR2>
 F;372-432/Domain: thrombospondin type 1 repeat homology <THR3>
 F;433-493/Domain: thrombospondin type 1 repeat homology <THR4>
 F;494-556/Domain: thrombospondin type 1 repeat homology <THR5>
 F;560-610/Domain: thrombospondin type 1 repeat homology <THR6>

Query Match 2.2%; Score 8; DB 2; Length 712;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 NAGGGAGG 240
 |||||
 Db 612 NAGGGAGG 619

RESULT 23
 S55344
 outer envelope membrane protein OBP75 precursor - garden pea
 N;Alternate names: Chloroplast import-associated channel IAP75
 C;Species: *Pisum sativum* (garden pea)
 C;Date: 14-Oct-1995 #sequence_revision 12-Apr-1996 #text_change 21-Jul-2000
 C;Accession: S55344; S55343; A55170; S51279
 R;Tranel, P.J.; Froehlich, J.; Goyal, A.; Keegstra, K.
 EMBO J. 14, 2436-2446, 1995
 A;Title: A component of the chloroplastic protein import apparatus is targeted to the outer envelope membrane protein OBP75 precursor.
 A;Reference number: S55343; MUID:95300776; PMID:7781598
 A;Accession: S55344
 A;Status: nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Residues: 1-809 <TRA>
 A;Cross-references: EMBL:X83767; NID:g633606; PIDN:CAA58720.1; PID:g633607

A;Molecule type: protein
 A;Residues: 136-150;185-196;360-367;532-545;642-660 <TRW>
 R;Schnebli, D.J.; Kessler, F.; Blobel, G.
 Science 266, 1007-1012, 1994
 A;Title: Isolation of components of the chloroplast protein import machinery.
 A;Reference number: A55170; MUID:95063931; PMID:7973649
 A;Accession: A55170
 A;Molecule type: mRNA
 A;Residues: 1-809 <SCH>
 A;Cross-references: GB:L36858; NID:g576506; PIDN:AAA53275.1; PID:g576507
 C;Genetics:
 A;Genome: nuclear
 C;Keywords: chloroplast; transmembrane protein
 F;1-131/Domain: transit peptide (chloroplast) #status predicted <TNP>
 F;132-809/Product: outer envelope membrane protein OBP75 #status predicted <MAT>

Query Match 2.2%; Score 8; DB 2; Length 809;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
 |||||
 Db 90 AGGGAGGG 97

RESULT 24
 T02812
 probable membrane protein L2759.4 [imported] - *Leishmania major* (strain Friedlin)
 C;Species: *Leishmania major*
 C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 19-May-2000
 C;Accession: H81457; T02812
 R;Myler, P.J.; Audleman, L.; deVos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magnus, C.
 Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
 A;Title: *Leishmania major* Friedlin chromosome 1 has an unusual distribution of protein

A;Reference number: AB1455; MUID:99178987; PMID:10077609
A;Accession: H81457
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-823 <PYL>
A;Cross-references: GB:AE001274; NID:g3264850; PIDN:AAC24635.1; PID:g2995588; GSPDB:GN00
A;Experimental source: strain MHOM/IL/81/Friedlin
C;Genetics:
A;Gene: L2759.4
A;Map position: 1

Query Match 2.2%; Score 8; DB 2; Length 823;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 ESTLEENK 63
|||||

DB 736 ESTLEENK 743

RESULT 25

brkA prtein - Bordetella pertussis

C;Species: Bordetella pertussis

C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 08-Oct-1999

C;Accession: I40329

R;Fernandez, R.C.; Weiss, A.A.

Infect. Immun. 62, 4727-4738, 1994

A;Title: Cloning and sequencing of a Bordetella pertussis serum resistance locus.

A;Reference number: I40328; MUID:95012680; PMID:7927748

A;Accession: I40329

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1010 <RES>

A;Cross-references: EMBL:U12276; NID:g562025; PIDN:AAA51646.1; PID:g562026

C;Genetics:

A;Gene: brkA

Query Match

Best Local Similarity 100.0%; Score 8; DB 2; Length 1010;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 QAPQPPVA 114
|||||

DB 43 QAPQPPVA 50

RESULT 26

T13709

diacylglycerol kinase (EC 2.7.1.107) - fruit fly (Drosophila melanogaster)

C;Species: Drosophila melanogaster

C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000

C;Accession: T13709

R;Masai, I.; Okazaki, A.; Hosoya, T.; Hotta, Y.

Proc. Natl. Acad. Sci. U.S.A. 90, 11157-11161, 1993

A;Title: Drosophila retinal degeneration A gene encodes an eye-specific diacylglycerol k

A;Reference number: Z17702; MUID:94068563; PMID:8248222

A;Accession: T13709

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-1454 <MAS>

A;Cross-references: EMBL:D17315; NID:d1007485; PID:d1004654; PIDN:BA04135.1

C;Genetics:

A;Gene: rdga

A;Cross-references: FlyBase:FBgn0003217

C;Keywords: phosphotransferase

Query Match

Best Local Similarity 100.0%; Score 8; DB 2; Length 1454;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
|||||

DB 764 AGGGAGGG 771

RESULT 27

T14603

hypotheical protein - Trypanosoma cruzi

C;Species: Trypanosoma cruzi

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 24-Sep-1999

C;Accession: T14603; T14634

R;Andersson, B.; Aslund, L.; Pettersson, U.

submitted to the EMBL Data Library, March 1998

A;Description: 93.4 kb of complete sequence from chromosome 3 of Trypanosoma cruzi.

A;Reference number: Z18159

A;Accession: T14603

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1718 <AND>

A;Cross-references: EMBL:AF052832; NID:g3063540; PID:g3063541; PIDN:AAC14077.1

A;Accession: T14634

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1718 <AN2>

A;Cross-references: EMBL:AF052833; NID:g3063554; PID:g3063567; PIDN:AAC14102.1

C;Genetics:

A;Map position: 3

Query Match

Best Local Similarity 100.0%; Score 8; DB 2; Length 1718;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
|||||

DB 543 AGGGAGGG 550

RESULT 28

D69426

surface layer protein B (slgB-2) homolog - Archaeoglobus fulgidus

C;Species: Archaeoglobus fulgidus

C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000

C;Accession: D69426

R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dode

.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes,

Smith, H.O.; Woese, C.R.; Venter, J.C.

A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archa

A;Reference number: A69250; MUID:98049343; PMID:9389475

A;Accession: D69426

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-2425 <KLE>

A;Cross-references: GB:AE001006; GB:AE000782; NID:g2689329; PIDN:AAB89834.1; PID:g2649

Query Match

Best Local Similarity 100.0%; Score 8; DB 2; Length 2425;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
|||||

DB 1144 AGGGAGGG 1151

RESULT 29

S77300

hypotheical protein slr1403 - Synecocystis sp. (strain PCC 6803)

C;Species: Synecocystis sp.

A;Variety: PCC 6803

C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999

C;Accession: S77300

R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, N

O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasu

DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S77300
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-3016 <KAN>
A;Cross-references: EMBL:D90907; GB:AB001339; NID:gl52618; PIDN:BAAL7634.1; PID:dl01836
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 2.2%; Score 8; DB 2; Length 3016;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGGG 241
|||||
Db 2629 AGGAGGG 2636

RESULT 30
A71429
hypothetical protein - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
A;Variety: Columbia
C;Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998
C;Accession: A71429
R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzengger, T.; Pohl, T.M.; Terry, N.; Giel
avanagh, T.; Hempel, S.; Kotter, F.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A;Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; And
C.; Chludzisz, N.
A;Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A;Reference number: A71400; MUID:9812113; PMID:9461215
A;Accession: A71429
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-42 <BEV>
A;Cross-references: GB:297340; NID:g2244950; PID:e327498; PID:g2244982
A;Genetics:
A;Map position: 4COP9-4G3845

Query Match 1.9%; Score 7; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
|||||
Db 21 GGGAGGG 27

RESULT 31
B49410
t-complex polypeptide 1 homolog (peak 6a fraction) - rabbit (fragments)
N;Alternate names: chaperonin homolog (peak 6a)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 21-Sep-1994 #sequence_revision 18-Nov-1994 #text_change 04-Sep-1998
C;Accession: B49410
R;Rommlaere, H.; Van Troys, M.; Gao, Y.; Melki, R.; Cowan, N.J.; Vandekerckhove, J.; Am
Proc. Natl. Acad. Sci. U.S.A. 90, 11975-11979, 1993
A;Title: Eukaryotic cytosolic chaperonin contains t-complex polypeptide 1 and seven rela
A;Reference number: A49410; MUID:94089752; PMID:7903455
A;Accession: B49410
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-50 <ROM>
A;Experimental source: reticulocyte
A;Note: sequence modified after extraction from NCBI backbone
A;Note: sequence extracted from NCBI backbone (NCBIP:141037)
C;Superfamily: molecular chaperone t-complex-type

Query Match 1.9%; Score 7; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 269 MLVELSK 275
|||||
Db 31 MLVELSK 37

RESULT 32
A39034
neurogranin - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Aug-1991 #sequence_revision 30-Aug-1991 #text_change 24-Nov-1999
C;Accession: A39034
R;Baudier, J.; Deloulme, J.C.; Van Dorsselaer, A.; Black, D.; Matthes, H.W.D.
J. Biol. Chem. 266, 229-237, 1991
A;Title: Purification and characterization of a brain-specific protein kinase C subst:
AP43) that corresponds to the protein kinase C phosphorylation site and the calmoduli
A;Reference number: A39034; MUID:91093129; PMID:1824695
A;Accession: A39034
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-76 <BAU>
C;Keywords: blocked amino end; phosphoprotein
F1/Modified site: blocked amino end (Met) (probably acetylated) #status experimental
F1/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 1.9%; Score 7; DB 2; Length 76;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 GGGAGGG 242
|||||
Db 67 GGGAGGG 73

RESULT 33
A57288
neurogranin - rat
N;Alternate names: gene RC3 protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 08-Oct-1999
C;Accession: A57288; S36982; S36881
R;Sato, T.; Xiao, D.M.; Li, H.; Huang, F.L.; Huang, K.P.
J. Biol. Chem. 270, 10314-10322, 1995
A;Title: Structure and regulation of the gene encoding the neuron-specific protein ki
A;Reference number: A57288; MUID:95247744; PMID:7730337
A;Accession: A57288
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-78 <SAT>
A;Cross-references: GB:U22062; NID:g924645; PIDN:AAA80223.1; PID:g924646
R;Watson, J.B.; Battenberg, E.F.; Wong, K.K.; Bloom, F.E.; Sutcliffe, J.G.
J. Neurosci. Res. 26, 397-408, 1990
A;Title: Subtractive cDNA cloning of RC3, a rodent cortex-enriched mRNA encoding a no
A;Reference number: S36982; MUID:91039409; PMID:2231781
A;Accession: S36982
A;Molecule type: mRNA
A;Residues: 1-78 <WAT>
A;Cross-references: EMBL:L09119; NID:g206594; PIDN:AAA42023.1; PID:g206595
R;Huang, K.P.; Huang, F.L.; Chen, H.C.
Arch. Biochem. Biophys. 305, 570-580, 1993
A;Title: Characterization of a 7.5-kDa protein kinase C substrate (RC3 protein, neuro:
A;Reference number: S36881; MUID:93384318; PMID:8080473
A;Accession: S36881
A;Molecule type: protein
A;Residues: 2-78 <HUA>

Query Match 1.9%; Score 7; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 GGAGGGP 242
 |||||
 Db 69 GGAGGGP 75

RESULT 34

I47043
 neurogranin RC3 [imported] - goat
 C:Species: Capra sp.
 C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 20-Jun-2000
 C:Accession: I47043
 R:Piosik, P.A.; van Groenigen, M.; Ponne, N.J.; Bolhuis, P.A.; Baas, F.
 Brain Res. Mol. Brain Res. 29, 119-130, 1995
 A:Title: RC3/neurogranin structure and expression in the caprine brain in relation to c
 A:Reference number: I47043; PMID:95287801; PMID:7539519
 A:Accession: I47043
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-78 <PIO>
 A:Cross-references: GB:S78295; NID:g994839; PIDN:AA834481.1; PID:g994841
 C:Genetics:
 A:Gene: RC3

Query Match 1.9%; Score 7; DB 2; Length 78;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 GGAGGGP 242
 |||||
 Db 69 GGAGGGP 75

RESULT 35

S49470
 hypothetical protein - Salmonella choleraesuis (fragment)
 C:Species: Salmonella choleraesuis
 C:Date: 20-Feb-1995 #sequence_revision 26-Jul-1996 #text_change 29-Sep-1999
 C:Accession: S49470
 R:Krause, M.; El-Gedaily, A.
 submitted to the EMBL Data Library, October 1994
 A:Reference number: S49470
 A:Accession: S49470
 A:Molecule type: DNA
 A:Residues: 1-94 <KRA>
 A:Cross-references: EMBL:X82129; NID:g558643; PIDN:CAA57639.1; PID:g558644
 A:Note: the source is designated as Salmonella enterica
 C:Superfamily: lipoprotein D

Query Match 1.9%; Score 7; DB 2; Length 94;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 GSTGTSS 87
 |||||
 Db 65 GSTGTSS 71

RESULT 36

F97355
 uncharacterized conserved protein, yuke/yfJA B. subtilis family CAC3713 [imported] - Clo
 C:Species: Clostridium acetobutylicum
 C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
 C:Accession: F97355
 R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
 J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4838, 2001
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
 A:Reference number: A96900; PMID:21359325; PMID:21359325
 A:Accession: F97355
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-96 <KUR>
 A:Cross-references: GB:AE001437; PIDN:AAK81633.1; PID:g15026818; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824
 C:Genetics:
 A:Gene: CAC3713

Query Match 1.9%; Score 7; DB 2; Length 96;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 ISVTPEE 325
 |||||
 Db 4 ISVTPEE 10

RESULT 37

S69752
 hypothetical protein YDR112w - Yeast (Saccharomyces cerevisiae)
 C:Species: Saccharomyces cerevisiae
 C:Date: 23-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 29-Oct-1999
 C:Accession: S69752
 R:Murphy, L.; Shore, L.; Harris, D.
 submitted to the EMBL Data Library, March 1995
 A:Reference number: S52671
 A:Accession: S69752
 A:Molecule type: DNA
 A:Residues: 1-102 <MUR>
 A:Cross-references: EMBL:Z48750; GSPDB:GN00004; MIPS:YDR112w
 C:Genetics:
 A:Gene: MIPS:YDR112w
 A:Map position: 4R

Query Match 1.9%; Score 7; DB 2; Length 102;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 SKGKTSG 81
 |||||
 Db 8 SKGKTSG 14

RESULT 38

T45500
 hypothetical protein [imported] - Thiobacillus ferrooxidans (fragment)
 C:Species: Thiobacillus ferrooxidans
 C:Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 21-Jul-2000
 C:Accession: T45500
 R:Oppon, J.C.; Sarnovsky, R.J.; Craig, N.L.; Rawlings, D.E.
 J. Bacteriol. 180, 3007-3012, 1998
 A:Title: A Tn7-like transposon is present in the glmUS region of the obligately chemo
 A:Reference number: Z22992; PMID:98269023; PMID:9603897
 A:Accession: T45500
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-102 <OPP>
 A:Cross-references: EMBL:AF032884; NID:g2653994; PIDN:AAC21668.1; PID:g2654001
 A:Experimental source: ATCC 33020

Query Match 1.9%; Score 7; DB 2; Length 102;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
 |||||
 Db 30 GGGAGGG 36

RESULT 39

G86252
 hypothetical protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002
 C:Accession: G86252
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpspiel, N.A.; Kaul, S.; White, O.; Alons
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K

ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 C;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A;Reference number: A86141; MUID:21016719; PMID:11130712
 A;Accession: G86252
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-108 <STO>
 A;Cross-references: GB:AE005172; NID:g3157948; PIDN:AAC17631.1; GSPDB:GN00141
 C;Genetics:
 A;Map position: 1
 C;Superfamily: glycine-rich RNA-binding protein; ribonucleoprotein repeat homology

Query Match 1.9%; Score 7; DB 2; Length 108;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
 |||||
 Db 97 GGGAGGG 103

RESULT 40
 C55993
 Hypothetical protein (pyp 3' region) - Ectothiorhodospira halophila (fragment)
 C;Species: Ectothiorhodospira halophila
 C;Date: 19-Jan-1996 #sequence_revision 19-Jan-1996 #text_change 08-Oct-1999
 A;Accession: C55993
 R;Baca, M.; Borgstahl, G.E.O.; Boissinot, M.; Burke, P.M.; Williams, D.R.; Slater, K.A.;
 Biochemistry 33, 14369-14377, 1994
 A;Title: Complete chemical structure of photoactive yellow protein: novel thioester-link
 A;Reference number: A55993; MUID:95072006; PMID:7981196
 A;Accession: C55993
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-112 <BAC>
 A;Cross-references: GB:U17017; NID:g502427; PIDN:AAA61736.1; PID:g602430

Query Match 1.9%; Score 7; DB 2; Length 112;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 TSGSTGT 85
 |||||
 Db 104 TSGSTGT 110

RESULT 41
 S58445
 Lipoprotein D - Salmonella typhimurium (fragment)
 C;Species: Salmonella typhimurium
 C;Date: 12-Feb-1998 #sequence_revision 15-May-1998 #text_change 29-Sep-1999
 A;Accession: S58445
 R;Robbe-Saule, V.; Coynault, C.; Norel, P.
 FEWS Microbiol. Lett. 126, 171-176, 1995
 A;Title: The live oral typhoid vaccine Ty21a is a rpoS mutant and is susceptible to vari
 A;Reference number: S58445; MUID:95220644; PMID:7705608
 A;Accession: S58445
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-119 <ROB>
 A;Cross-references: EMBL:X81641; NID:g695757; PIDN:CAA57297.1; PID:g695758
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1994
 C;Genetics:
 A;Gene: nlpD
 C;Superfamily: lipoprotein D

Query Match 1.9%; Score 7; DB 2; Length 119;

Best Local Similarity 100.0%; Pred. No. 36;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 GSGTSS 87
 |||||
 Db 90 GSGTSS 96

RESULT 42
 DB3723
 Hypothetical protein BH0588 [imported] - Bacillus halodurans (strain C-125)
 C;Species: Bacillus halodurans
 C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C;Accession: DB3723
 R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hi
 Nucleic Acids Res. 28, 4317-4331, 2000
 A;Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* ar
 A;Reference number: AB3650; MUID:20512582; PMID:11058132
 A;Accession: DB3723
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-119 <STO>
 A;Cross-references: GB:AP001509; GB:BA000004; NID:g10173176; PIDN:BA04307.1; GSPDB:G
 A;Experimental source: strain C-125
 C;Genetics:
 A;Gene: BH0588
 C;Superfamily: Escherichia coli hypothetical 14K protein (cysG-trpS region)

Query Match 1.9%; Score 7; DB 2; Length 119;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 177 BRAVEYL 183
 |||||
 Db 106 BRAVEYL 112

RESULT 43
 A24272
 Ig heavy chain precursor V region (RVH832) - rabbit (fragment)
 C;Species: Oryctolagus cuniculus (domestic rabbit)
 C;Date: 08-Aug-1987 #sequence_revision 08-Aug-1987 #text_change 21-Jan-2000
 C;Accession: A24272
 R;McCormack, W.T.; Laster, S.M.; Marzluff, W.F.; Roux, K.H.
 Nucleic Acids Res. 13, 7041-7054, 1985
 A;Title: Dynamic gene interactions in the evolution of rabbit Vh genes: a four codon
 A;Reference number: A24272; MUID:86041895; PMID:2997735
 A;Accession: A24272
 A;Molecule type: DNA
 A;Residues: 1-121 <MCC>
 A;Cross-references: GB:X03054; NID:g1597; PIDN:CAA26862.1; PID:g1598
 A;Note: this sequence was determined from the germline gene
 C;Genetics:
 A;Introns: 15/3
 C;Superfamily: immunoglobulin V region; immunoglobulin homology
 C;Keywords: heterotrimer; immunoglobulin
 P;37-121/Domain: immunoglobulin homology <IMM>

Query Match 1.9%; Score 7; DB 2; Length 121;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
 |||||
 Db 26 GGGAGGG 32

RESULT 44
 T36107
 Probable serine/arginine rich protein - Streptomyces coelicolor
 C;Species: Streptomyces coelicolor
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C;Accession: T36107

R;Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, April 1999
 A;Reference number: Z21597
 A;Accession: T36107
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-172 <MUR>
 A;Cross-references: EMBL:AL049707; PIDN:CAB41273.1; GSPDB:GN00070; SCOEDB:SCE15.04
 A;Experimental source: strain A3(2)
 C;Genetics:
 A;Gene: SCOEDB:SCE15.04

Query Match 1.9%; Score 7; DB 2; Length 172;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 GGGAGGG 241
 |||||
 Db 51 GGGAGGG 57

RESULT 45
 C82243
 conserved hypothetical protein VC1075 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
 C;Species: Vibrio cholerae
 C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C;Accession: C82243
 R;Heidelberg, J.F.; Easen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, P.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: C82243

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-180 <HEI>

A;Cross-references: GB:AR004189; GB:AE003852; NID:G9655541; PIDN:AAF94234.1; GSPDB:GN001

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC1075

A;Map position: 1

C;Superfamily: conserved hypothetical protein ybhB

Query Match 1.9%; Score 7; DB 2; Length 180;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 228 PQGASNA 234
 |||||
 Db 151 PQGASNA 157

RESULT 46
 I59203
 gene HOX2.8 protein - human (fragment)
 C;Species: Homo sapiens (man)
 C;Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 15-Oct-1999
 C;Accession: I59203
 R;Lavie, D.; Ducksworth, J.; Eves, E.; Gomes, G.; Keller, M.; Heller, P.; DeSimone, J. Proc. Natl. Acad. Sci. U.S.A. 88, 7318-7322, 1991
 A;Title: A homeodomain protein binds to gamma-globin gene regulatory sequences.
 A;Reference number: I59203; MUID:913334457; PMID:1871139

A;Accession: I59203

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-182 <RES>

A;Cross-references: GB:M73999; NID:G184283; PIDN:AAB00778.1; PID:G184284

C;Superfamily: unassigned homeobox proteins; homeobox homology

C;Keywords: DNA binding; homeobox; nucleus; transcription regulation

F:58-114/Domain: homeobox homology <HOX>

Query Match 1.9%; Score 7; DB 2; Length 182;

Best Local Similarity 100.0%; Pred. No. 53;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 234 AGGAGGG 240
 |||||
 Db 162 AGGAGGG 168

RESULT 47

AF0830

probable membrane protein STY2838 [imported] - Salmonella enterica subsp. enterica ser C;Species: Salmonella enterica subsp. enterica serovar Typhi

A;Note: This species has also been called Salmonella typhi

C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C;Accession: AF0830

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farra, S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.

A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica ser

A;Reference number: AB0502; MUID:21534947; PMID:11677608

A;Accession: AF0830

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-195 <PAR>

A;Cross-references: GB:AL513382; PIDN:CAD02794.1; PID:G16503803; GSPDB:GN00176

C;Genetics:

A;Gene: STY2838

Query Match 1.9%; Score 7; DB 2; Length 195;

Best Local Similarity 100.0%; Pred. No. 57;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 68 GFLVWML 74
 |||||
 Db 49 GFLVWML 55

RESULT 48

T00799

hypothetical protein At2g32690 [imported] - Arabidopsis thaliana

N;Alternate names: hypothetical protein F24L7.17

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 16-Feb-2001

C;Accession: T00799; C84736

R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Ka

submitted to the EMBL Data Library, February 1998

A;Description: Arabidopsis thaliana chromosome II BAC F24L7 genomic sequence.

A;Reference number: Z14204

A;Accession: T00799

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-201 <ROU>

A;Cross-references: EMBL:AC003974; NID:G2914688; PIDN:AAC04494.1; PID:G2914704

A;Experimental source: cultivar Columbia

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon,

euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: C84736

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-201 <STO>

A;Cross-references: GB:AE002093; NID:G2914704; PIDN:AAC04494.1; GSPDB:GN00139

C;Genetics:

A;Gene: At2g32690; F24L7.17

A;Map position: 2

C;Superfamily: glycine-rich cell wall structural protein 1

Query Match 1.9%; Score 7; DB 2; Length 201;

```

Best Local Similarity 100.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
   |||||
Db 138 GGGAGGG 144

RESULT 49
G01204
twist protein homolog - human
C:Species: Homo sapiens (man)
C>Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 08-Oct-1999
C:Accession: G01204; JC6330
R:Cristofalo, V.J.
submitted to the EMBL Data Library, September 1995
A:Reference number: G06124
A:Accession: G01204
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-201 <CR1>
A:Cross-references: EMBL:X91662; NID:g999455; PIDN:CAA62850.1; PID:g999456
R:Wang, S.M.; Coljee, V.W.; Pignolo, R.J.; Rotenberg, M.O.; Cristofalo, V.J.; Sierra, F.
Gene 187, 83-92, 1997
A:Title: Cloning of the human twist gene: Its expression is retained in adult mesodermal
A:Reference number: JC6330; MUID:97225800; PMID:9073070
A:Accession: JC6330
A:Molecule type: mRNA
A:Residues: 1-52, 'A', 54-201 <WAN>
A:Cross-references: GB:X91662; NID:g999455
C:Comment: This protein is a basic helix-loop-helix DNA-binding transcription factor the
C:Genetics:
A:Gene: GDB:TWIST
A:Cross-references: GDB:6278991; OMIM:601622
A:Map position: 7p21-7p21
A:Keywords: DNA binding; transcription factor
F:108-120/Domain: DNA-binding #status Predicted <DNB>
F:121-201/Region: helix-loop-helix #status predicted

Query Match 1.9%; Score 7; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
   |||||
Db 83 GGGAGGG 89

RESULT 50
I53066
gene M-twist protein - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999
C:Accession: I53066; I66795
R:Wolf, C.; Thisee, C.; Stoetzel, C.; Thisee, B.; Gerlinger, P.; Perrin-Schmitt, F.
Dev. Biol. 143, 363-373, 1991
A:Title: The M-twist gene of Mus is expressed in subsets of mesodermal cells and is clon
A:Reference number: I53066; MUID:91122450; PMID:1840517
A:Accession: I53066
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-206 <RES>
A:Cross-references: GB:M63649; NID:g202243; PIDN:AAA40514.1; PID:g202244
A:Accession: I66795
A:Molecule type: mRNA
A:Residues: 1-35, 'R', 37-90, 'P', 92-206 <RE2>
A:Cross-references: GB:M63650; NID:g202245; PIDN:AAA40515.1; PID:g202246
C:Genetics:
A:Gene: M-twist

Query Match 1.9%; Score 7; DB 2; Length 206;
Best Local Similarity 100.0%; Pred. No. 59;

```

```

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
   |||||
Db 84 GGGAGGG 90

RESULT 51
T07854
germin-like protein (clone BnC4) - rape
C:Species: Brassica napus (rape)
C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 08-Oct-1999
C:Accession: T07854
R:Saenz-Vasquez, J.
submitted to the EMBL Data Library, February 1995
A:Reference number: Z16177
A:Accession: T07854
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-207 <SAE>
A:Cross-references: EMBL:U21743; NID:g914910; PIDN:AAA86365.1; PID:g914911
A:Experimental source: cv. Samourai; 5 days old seedlings
C:Superfamily: germin

Query Match 1.9%; Score 7; DB 2; Length 207;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TVKTLKG 10
   |||||
Db 195 TVKTLKG 201

RESULT 52
H69037
conserved hypothetical protein MTH1284 - Methanobacterium thermoautotrophicum (strain
C:Species: Methanobacterium thermoautotrophicum
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: H69037
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني,
K. S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fur
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: H69037
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-209 <MTH>
A:Cross-references: GB:AE000893; GB:AE000666; NID:g2622375; PIDN:AAB85766.1; PID:g2622
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH1284

Query Match 1.9%; Score 7; DB 2; Length 209;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 GPLDPLR 247
   |||||
Db 104 GPLDPLR 110

RESULT 53
CB7542
hypothetical protein CC2364 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: CB7542
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kol
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

```

A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
 A:Reference number: AB7249; MUID:21173698; PMID:11259647
 A:Accession: C87542
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-215 <STO>
 A:Cross-references: GB:AE005673; NID:gl3423895; PIDN:AAK24335.1; GSPDB:GN00148
 C:Genetics:
 A:Gene: CC2364

Query Match 1.9%; Score 7; DB 2; Length 215;
 Best Local Similarity 100.0%; Pred. No. 62;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 QPEGLPA 126
 |||||
 Db 124 QPEGLPA 130

RESULT 54

T48482
 hypothetical protein T28J14.50 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jun-2000
 R:Accession: T48482
 R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: T24493
 A:Accession: T48482
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-216 <BEV>
 A:Cross-references: EMBL:AL163652
 A:Experimental source: cultivar Columbia; BAC clone T28J14
 C:Genetics:
 A:Map position: 5
 A:Note: T28J14.50
 C:Superfamily: Arabidopsis thaliana hypothetical protein T28J14.50

Query Match 1.9%; Score 7; DB 2; Length 216;
 Best Local Similarity 100.0%; Pred. No. 62;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 324 EEQEAIG 330
 |||||
 Db 181 EEQEAIG 187

RESULT 55

AF3273
 ribosomal protein L11 methyltransferase (EC 2.1.1.-) [imported] - *Brucella melitensis* (e
 C:Species: *Brucella melitensis*
 C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
 R:Delevecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujter, C.; Los, T.; Ivanova,
 ; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
 A:Reference number: AD3252; PMID:11756688
 A:Accession: AF3273
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-218 <KUR>
 A:Cross-references: GB:AE008917; PIDN:AAL51353.1; PID:gl7982052; GSPDB:GN00190
 A:Experimental source: strain 16N.
 C:Genetics:
 A:Gene: BMEI0171
 A:Map position: 1
 C:Keywords: methyltransferase

Query Match 1.9%; Score 7; DB 2; Length 218;
 Best Local Similarity 100.0%; Pred. No. 63;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 AISVTPE 324
 |||||
 Db 128 AISVTPE 134

RESULT 56

A53143
 testis-determining gene/SRY homolog - dunart (*Sminthopsis macroura*) (fragment)
 C:Species: *Sminthopsis macroura*
 C:Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 24-Sep-1999
 R:Accession: A53143
 R:Poster, J.W.; Graves, J.A.
 Proc. Natl. Acad. Sci. U.S.A. 91, 1927-1931, 1994
 A:Title: An SRY-related sequence on the marsupial X chromosome: implications for the e
 A:Reference number: A53143; MUID:94173941; PMID:8127908
 A:Accession: A53143
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-224 <FOS>
 A:Cross-references: GB:S69429; NID:gs45827; PIDN:AAB30154.1; PID:gs45828
 A:Note: sequence inconsistent with nucleotide translation
 A:Note: sequence extracted from NCBI backbone (NCBIN:145890, NCBIIP:145892)
 C:Superfamily: unassigned HMG box proteins; HMG box homology
 F:52-127/Domain: HMG box homology <HMG>

Query Match 1.9%; Score 7; DB 2; Length 224;
 Best Local Similarity 100.0%; Pred. No. 64;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
 |||||
 Db 41 GGGAGGG 47

RESULT 57

A98166
 hypothetical protein AGR_L_563 [imported] - *Agrobacterium tumefaciens* (strain C58, Cer
 C:Species: *Agrobacterium tumefaciens*
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
 R:Accession: A98166
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldma
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markeiz, B.
 Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium t*
 A:Reference number: A97359; MUID:21608551; PMID:11743194
 A:Accession: A98166
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-228 <KUR>
 A:Cross-references: GB:AE007870; PIDN:AAK88851.1; PID:gl5158615; GSPDB:GN00170
 C:Genetics:
 A:Gene: AGR_L_563
 A:Map position: linear chromosome

Query Match 1.9%; Score 7; DB 2; Length 228;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 DNAASNL 139
 |||||
 Db 154 DNAASNL 160

RESULT 58

AF3121
 hypothetical protein Atu4594 [imported] - *Agrobacterium tumefaciens* (strain C58, Dupon
 C:Species: *Agrobacterium tumefaciens*
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 R:Accession: AF3121
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCle
 ; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AF3121
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-228 <RUR>
A:Cross-references: GB:AE008689; PIDN:AAL45388.1; PID:G17743085; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atcu4594
A:Map position: linear chromosome

Query Match 1.9%; Score 7; DB 2; Length 228;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 DNAASNL 139
Db 154 DNAASNL 160

RESULT 59
T07358
ribosomal protein S3 - *Chlorella vulgaris* chloroplast
C:Species: chloroplast *Chlorella vulgaris*
C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 21-Jul-2000
C:Accession: T07358
R:Nakauchi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Nakag Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997
A:Title: Complete nucleotide sequence of the chloroplast genome from the green alga *Chlo*
A:Reference number: Z15985; MUID:97303241; PMID:9159184
A:Accession: T07358
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-231 <WAK>
A:Cross-references: EMBL:AB001684; NID:G2224352; PIDN:BAA58006.1; PID:G2224522
C:Genetics:
A:Gene: rps3
A:Genome: chloroplast
C:Superfamily: Escherichia coli ribosomal protein S3
C:Keywords: chloroplast; ribosome

Query Match 1.9%; Score 7; DB 2; Length 231;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 VQRALRA 171
Db 154 VQRALRA 160

RESULT 60
G95989
hypothetical glycine-rich protein [imported] - *Sinorhizobium meliloti* (strain 1021) mega
C:Species: *Sinorhizobium meliloti*
C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: G95989
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernar Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb pSymb megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: G95989
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-234 <RUR>
A:Cross-references: GB:ALS1985; PIDN:CAC49583.1; PID:G15141070; GSPDB:GN00167
A:Experimental source: strain 1021, megaplasmid pSymb
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, P.; Barloy-Hubler, pella, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kliss, E.; Komp, C.; Lelau hebaul, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMB20889
A:Genome: plasmid

Query Match 1.9%; Score 7; DB 2; Length 234;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGG 240
Db 176 AGGGAGG 182

RESULT 61
I46082
CD8 alpha-chain - cat
C:Species: *Felis silvestris catus* (domestic cat)
C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 23-Jul-1999
C:Accession: I46082
R:Pecoraro, M.; Kawaguchi, Y.; Miyazawa, T.; Norimine, J.; Maeda, K.; Toyosaki, T.; T Immunology 81, 127-131, 1994
A:Title: Isolation, sequence and expression of a cDNA encoding the alpha-chain of the
A:Reference number: I46082; MUID:94178799; PMID:8132208
A:Accession: I46082
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-239 <PEC>
A:Cross-references: GB:D16536; NID:G485385; PIDN:BAA03973.1; PID:G485386
C:Superfamily: immunoglobulin V region, immunoglobulin homology

Query Match 1.9%; Score 7; DB 2; Length 239;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 APITTSQ 120
Db 153 APITTSQ 159

RESULT 62
T46968
flavoprotein [imported] - *Paracoccus denitrificans* (fragment)
C:Species: *Paracoccus denitrificans*
C>Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 21-Jul-2000
C:Accession: T46968
R:Wodara, C.; Bardischewsky, F.; Friedrich, C.G.
J. Bacteriol. 179, 5014-5023, 1997
A:Title: Cloning and characterization of sulfite dehydrogenase, two c-type cytochromes sulfur oxidation.
A:Reference number: Z24324; MUID:97405897; PMID:9260941
A:Accession: T46968
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-247 <WOD>
A:Cross-references: EMBL:X79242; NID:G2253074; PIDN:CAA55826.1; PID:G2222781
A:Experimental source: strain GB17
C:Genetics:
A:Gene: soxP

Query Match 1.9%; Score 7; DB 2; Length 247;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGG 241
Db 36 GGGAGG 42

RESULT 63

T33230
 hypothetical protein T10B5.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T33230
 R:Giesel, C.; Bradshaw, H.; O'Brien, D.
 A:Submitted to the EMBL Data Library, May 1998
 A:Description: The sequence of C. elegans cosmid T10B5.
 A:Reference number: Z21305
 A:Accession: T33230
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-248 <GEI>
 A:Cross-references: EMBL:AF067947; PIDN:AAC19227.1; GSPDB:GN00023; CESP:T10B5.3
 A:Experimental source: strain Bristol N2; clone T10B5
 C:Genetics:
 A:Gene: CESP:T10B5.3
 A:Map position: 5
 A:Introns: 65/2; 141/2

Query Match 1.9%; Score 7; DB 2; Length 248;
 Best Local Similarity 100.0%; Pred. No. 71;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 76 KGKTSGS 82

Db 72 KGKTSGS 78

RESULT 64

B33144
 homeotic protein Ultrabithorax - house fly (fragment)
 C:Species: Musca domestica (house fly)
 C>Date: 30-Nov-1990 #sequence_revision 20-Sep-1991 #text_change 18-Jun-1993
 C:Accession: B33144
 R:Wilde, C.D.; Akam, M.
 EMBO J. 6, 1393-1401, 1987
 A:Title: Conserved sequence elements in the 5' region of the Ultrabithorax transcription
 A:Reference number: A33144
 A:Accession: B33144
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-249 <WIL>
 A:Cross-references: GB:X05178; GB:Y00066

Query Match 1.9%; Score 7; DB 2; Length 249;
 Best Local Similarity 100.0%; Pred. No. 71;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 GGGAGGG 241

Db 108 GGGAGGG 114

RESULT 65

B84777
 hypothetical protein At2g36120 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: B84777
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon, L.;
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: B84777
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-255 <STO>
 A:Cross-references: GB:AE002093; NID:G4678224; PIDN:AAD26969.1; GSPDB:GN00139

Query Match 1.9%; Score 7; DB 2; Length 268;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

C:Genetics:

A:Gene: At2g36120
 A:Map position: 2

C:Superfamily: collagen alpha 1(V) chain; fibrillar collagen carboxyl-terminal homolog

Query Match 1.9%; Score 7; DB 2; Length 255;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 GGGAGGG 241

Db 89 GGGAGGG 95

RESULT 66

CIRBL
 calpain (EC 3.4.22.17) small chain - rabbit
 A:Alternate names: calcium-activated neutral proteinase (CANP); calpain light chain; c
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C>Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999
 C:Accession: A24816
 R:Emori, Y.; Kawasaki, H.; Imajoh, S.; Kawashima, S.; Suzuki, K.
 J. Biol. Chem. 261, 9472-9476, 1986
 A:Title: Isolation and sequence analysis of cDNA clones for the small subunit of rabbi
 A:Reference number: A24816; MUID:86250903; PMID:3013892
 A:Accession: A24816
 A:Molecule type: mRNA
 A:Residues: 1-266 <EMO>
 A:Cross-references: GB:M13364; NID:G164875; PIDN:AA81565.1; PID:G164876
 C:Complex: heterodimer of L (large) and S (small) chains
 C:Function:

A:Description: catalyzes the hydrolysis of peptides
 A:Note: cleaves preferentially after tyrosine, methionine, or arginine residues and be
 C:Superfamily: calpain small chain; calmodulin repeat homology
 C:Keywords: calcium binding; cysteine proteinase; duplication; EF hand; heterodimer; h
 F:1-54/Domain: glycine-rich <GLY>
 F:94-125/Domain: calmodulin repeat homology <EF1>
 F:137-169/Domain: calmodulin repeat homology <EF2>
 F:170-199/Domain: calmodulin repeat homology <EF3>
 F:202-234/Domain: calmodulin repeat homology <EF4>
 F:235-266/Domain: calmodulin repeat homology <EF5>

Query Match 1.9%; Score 7; DB 1; Length 266;
 Best Local Similarity 100.0%; Pred. No. 75;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 GGGAGGG 241

Db 47 GGGAGGG 53

RESULT 67

S09860
 hypothetical protein UL112 - human cytomegalovirus (strain AD169) (fragment)
 C:Species: human cytomegalovirus, human herpesvirus 5
 A:Note: host Homo sapiens (man)
 C>Date: 07-Sep-1990 #sequence_revision 31-Dec-1991 #text_change 27-Jan-1995
 C:Accession: S09860
 R:Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.
 M.; Barrell, B.G.
 Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
 A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovir
 A:Reference number: S09749; MUID:90269039; PMID:2161319
 A:Accession: S09860
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-268 <CHE>
 A:Cross-references: EMBL:X17403
 A:Note: this sequence was submitted to the EMBL Data Library, December 1989

Query Match 1.9%; Score 7; DB 2; Length 268;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
|||||
Db 206 GGGAGGG 212

RESULT 68

G83023
3',5'-cyclic-nucleotide phosphodiesterase (BC 3.1.4.17) cpdA-type PA4969 [similarity] -
N;Alternate names: icc protein
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 24-Aug-2001
C;Accession: G83023
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: G83023
A;Molecule type: DNA
A;Residues: 1-272 <STO>
A;Cross-references: GB:AE004910; GB:AE004091; NID:g9951251; PIDN:AA08354.1; GSPDB:GN001186
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA4969
C;Superfamily: 3',5'-cyclic-nucleotide phosphodiesterase cpdA; 3',5'-cyclic-nucleotide phosphodiesterase
C;Keywords: iron; metalloprotein; phosphoric diester hydrolase
F;15-202/Domain: 3',5'-cyclic-nucleotide phosphodiesterase cpdA homology <CPDA>
F;15-95/Domain: phosphodiesterase core homology <PEC>

Query Match 1.9%; Score 7; DB 1; Length 272;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 HSNTPAT 96
|||||
Db 4 HSNTPAT 10

RESULT 69

A83583
probable biotin synthesis protein BioC PA0503 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: A83583
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: A83583
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-274 <STO>
A;Cross-references: GB:AE004487; GB:AE004091; NID:g9946361; PIDN:AA03892.1; GSPDB:GN001186
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA0503

Query Match 1.9%; Score 7; DB 2; Length 274;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 QPEGLPA 126
|||||
Db 255 QPEGLPA 261

RESULT 70

AH2885
hypothetical protein dppD [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AH2885
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, S.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClure, G.; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: A82577; MUID:21608550; PMID:11743193
A;Accession: AH2885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-276 <KUR>
A;Cross-references: GB:AE008688; PIDN:AA43502.1; PID:g17741010; GSPDB:GN00186
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: dppD
A;Map position: circular chromosome

Query Match 1.9%; Score 7; DB 2; Length 276;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 ILRLIEE 286
|||||
Db 190 ILRLIEE 196

RESULT 71

E97661
dipeptide transporter dppD homolog (AB026907) [imported] - Agrobacterium tumefaciens (strain C58)
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C;Accession: E97661
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorllo, B.; Goldma, A.; Liu, F.; Mollam, C.; Allinger, D.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, I.
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: E97661
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-276 <KUR>
A;Cross-references: GB:AE007869; PIDN:AAK88246.1; PID:g15157704; GSPDB:GN00169
C;Genetics:
A;Gene: AGR_C_4569
A;Map position: circular chromosome

Query Match 1.9%; Score 7; DB 2; Length 276;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 ILRLIEE 286
|||||
Db 190 ILRLIEE 196

RESULT 72

S74256
homeotic protein six3-alpha - mouse
C;Species: Mus musculus (house mouse)
C;Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 02-Jul-1998
C;Accession: S74256
R;Kawakami, K.; Ohto, H.; Takizawa, T.; Saito, T.
FEBS Lett. 393, 259-263, 1996
A;Title: Identification and expression of six family genes in mouse retina.
A;Reference number: S74253; MUID:96409319; PMID:8814301
A;Accession: S74256
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-284 <KAW>

A:Cross-references: EMBL:D83147
 A:Experimental source: retina; strain BALB/c
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulation
 F:208-264/Domain: homeobox homology <HGX>

Query Match 1.9%; Score 7; DB 2; Length 284;
 Best Local Similarity 100.0%; Pred. No. 80;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241

Db 42 GGGAGGG 48

RESULT 73

T23416

hypothetical protein K07F5.11 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000

C:Accession: T23416

R:Hembry, C.

submitted to the EMBL Data Library, March 1996

A:Reference number: Z19738

A:Accession: T23416

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-290 <WIL>

A:Cross-references: EMBL:Z70284; PIDN:CAA94280.1; GSPDB:GN00022; CESP:K07F5.11

A:Experimental source: clone K07F5

C:Genetics:

A:Gene: CESP:K07F5.11

A:Map position: 4

A:Introns: 89/1; 257/1

C:Superfamily: Phaseolus glycine-rich cell wall protein 1.8

Query Match 1.9%; Score 7; DB 2; Length 290;
 Best Local Similarity 100.0%; Pred. No. 82;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241

Db 275 GGGAGGG 281

RESULT 74

S31415

glycine-rich protein GRP22 - rape.

C:Species: Brassica napus (rape)

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: S31415

R:Bergeron, D.; Boivin, R.; Raszczynski, C.L.; Bellemare, G.

submitted to the EMBL Data Library, August 1992

A:Description: Characterization and expression of a gene family encoding glycine-rich pr

A:Reference number: S31415

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-231 <BER>

A:Cross-references: EMBL:Z15045; NID:G17820; PIDN:CAA78762.1; PID:G17821

C:Superfamily: Phaseolus glycine-rich cell wall protein 1.8

Query Match 1.9%; Score 7; DB 1; Length 291;
 Best Local Similarity 100.0%; Pred. No. 82;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241

Db 85 GGGAGGG 91

RESULT 75

T00503

probable MYB family transcription factor [imported] - Arabidopsis thaliana
 N:Alternate names: myb-related protein homolog T20D16.8
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 23-Mar-2001
 C:Accession: T00503; H84622
 R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crooby, M.L.; Brandon, R.C.; Sykes, S.M.; Ke
 submitted to the EMBL Data Library, November 1997
 A:Description: Arabidopsis thaliana chromosome II BAC T20D16 genomic sequence.
 A:Reference number: Z14164
 A:Accession: T00503

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-309 <ROU>

A:Cross-references: EMBL:AC002391; NID:G2642427; PIDN:AAB87103.1; PID:G2642435

A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon,

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: H84622

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-309 <STO>

A:Cross-references: GB:AE002093; NID:G2642435; PIDN:AAB87103.1; GSPDB:GN00139

C:Genetics:

A:Gene: T20D16.8; At2g23290

A:Map position: 2

C:Superfamily: Arabidopsis myb-related 33.2K protein; myb DNA-binding repeat homology

F:8-59/Domain: myb DNA-binding repeat homology <MYB>

F:60-110/Domain: myb DNA-binding repeat homology <MYB2>

Query Match 1.9%; Score 7; DB 2; Length 309;
 Best Local Similarity 100.0%; Pred. No. 86;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241

Db 277 GGGAGGG 283

Search completed: December 17, 2003, 06:36:01

Job time : 21.6624 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 06:21:26 ; Search time 10.4735 Seconds
(without alignments)
1652.347 Million cell updates/sec

Title: US-09-805-550-4
Perfect score: 368
Sequence: 1 MKLTVKTLKGTHTFEIRVQPN.....CDRNEELAANYLLEHAGEED 368

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	2.4	368	1 RH23 SCHPO	O74803 schizosacch
2	8	2.2	331	1 SHX2 HUMAN	O60902 homo sapten
3	8	2.2	331	1 SHX2 MOUSE	P70390 mus musculus
4	8	2.2	349	1 GRP ARATH	P27483 arabidopsis
5	8	2.2	384	1 GRP PETHY	P09789 petunia hyb
6	8	2.2	456	1 FXD1 MOUSE	O61345 mus musculus
7	8	2.2	465	1 GRP2 PHAVU	P10496 phaseolus v
8	8	2.2	497	1 IRBP2 HUMAN	O60548 homo sapten
9	8	2.2	610	1 IRBP MOUSE	P49194 mus musculus
10	8	2.2	631	1 YC1Q ECOLI	P45848 escherichia
11	8	2.2	828	1 CADM HUMAN	O90199 homo sapten
12	8	2.2	1454	1 KDGE DROME	O09103 drosophila
13	7	1.9	76	1 NEUG BOVIN	P35722 bos taurus
14	7	1.9	78	1 NEUG CAPHI	P54877 capra hircu
15	7	1.9	78	1 NEUG HUMAN	O92686 homo sapten
16	7	1.9	78	1 NEUG RAT	O04940 rattus norv
17	7	1.9	96	1 Y1B3 CLOAB	P34159 clostridium
18	7	1.9	97	1 NLFD YEREN	P47764 yersinia en
19	7	1.9	110	1 PER DROER	O26288 drosophila
20	7	1.9	141	1 HBB DASAK	P56592 dasyatis ak
21	7	1.9	202	1 TWS1 HUMAN	O15672 homo sapten
22	7	1.9	206	1 TWS1 MOUSE	P26687 mus musculus
23	7	1.9	207	1 GLP1 BRANA	P46271 brassica na
24	7	1.9	231	1 RRS3 CHLVU	P56365 chlorella v
25	7	1.9	239	1 CDBA FELCA	P41688 felis ailve
26	7	1.9	248	1 HMXK MUSDO	O25451 musca domes
27	7	1.9	264	1 GSHI HUMAN	O90482 homo sapten
28	7	1.9	266	1 CANS RABIT	P06813 corytolagus
29	7	1.9	268	1 EP34 HCMVA	P16768 human cytom
30	7	1.9	303	1 CDK4 FIG	P79432 sus scrofa
31	7	1.9	306	1 RALY HUMAN	O90km9 homo sapten
32	7	1.9	320	1 DDL XYLFA	O9pf79 xylella fae
33	7	1.9	323	1 HXDB MOUSE	P233013 mus musculus

ALIGNMENTS

RESULT 1

34	7	1.9	329	1	A32C DROME	O46203 drosophila
35	7	1.9	329	1	TAL MOUSE	P22091 mus musculus
36	7	1.9	332	1	SIX3 HUMAN	O95343 homo sapien
37	7	1.9	333	1	SIX3 MOUSE	O62333 mus musculus
38	7	1.9	339	1	SYPA_CLOAB	Q97gk9 clostridium
39	7	1.9	342	1	ROAL_SCHAM	P21522 schistocercu
40	7	1.9	367	1	BET3 MESAU	O09029 mesocricetu
41	7	1.9	373	1	NLPD SALT1	O56131 salmomella
42	7	1.9	375	1	GLK1 TRIVA	O9gtw9 trichomonas
43	7	1.9	376	1	DNAB_LISIN	O92bn9 listeria in
44	7	1.9	377	1	DNAB_LISMO	O985a3 listeria mo
45	7	1.9	377	1	NLPD_SALDU	P39700 salmonella
46	7	1.9	377	1	NLPD SALT1	P40827 salmonella
47	7	1.9	377	1	RNG1 HUMAN	O06587 homo sapien
48	7	1.9	378	1	DNJ2_STRCO	O9rdd7 streptomyce
49	7	1.9	379	1	NLPD_ECOLI	P33648 escherichia
50	7	1.9	383	1	TCF7_HUMAN	P36402 homo sapien
51	7	1.9	385	1	RO32_XENLA	P51992 xenopus lae
52	7	1.9	389	1	DNAB_HALCU	O34135 halobacteri
53	7	1.9	391	1	DNAB_HALN1	O9hry3 halobacteri
54	7	1.9	391	1	PCL_ECTHA	P42516 ectothiorho
55	7	1.9	393	1	HXA_A HUMAN	P31260 homo sapien
56	7	1.9	406	1	HEM1_AQUAE	O67314 aquifex aeo
57	7	1.9	413	1	YB44_MYCPN	P75142 mycoplasma
58	7	1.9	416	1	SO DROME	Q27350 drosophila
59	7	1.9	417	1	PCL_RHOCA	O69140 rhodobacter
60	7	1.9	421	1	COG8 RAT	O8vhw5 rattus norv
61	7	1.9	423	1	COG8 MOUSE	Q8vhw2 mus musculus
62	7	1.9	423	1	MKR2_SERQU	Q9dd48 seriola qui
63	7	1.9	425	1	COG8_HUMAN	Q8wx85 homo sapien
64	7	1.9	429	1	SR54_THEAQ	O07347 thermus aqu
65	7	1.9	433	1	ELT2_CABEL	Q10655 caenorhabdi
66	7	1.9	437	1	KLCM MOUSE	P08730 mus musculus
67	7	1.9	440	1	FXGA_CHICK	O98937 gallus gall
68	7	1.9	440	1	SOX4_MOUSE	Q08831 mus musculus
69	7	1.9	441	1	SK11_HUMAN	P35716 homo sapien
70	7	1.9	443	1	COBB_METJA	Q58816 methanococc
71	7	1.9	454	1	GLNA_HALVO	P43386 halobacteri
72	7	1.9	465	1	FXD1_HUMAN	Q16676 homo sapien
73	7	1.9	465	1	FXD3_MOUSE	Q61060 mus musculus
74	7	1.9	467	1	SYC_THETN	Q8r7t3 thermoanar
75	7	1.9	486	1	KIC0_XENLA	P02537 xenopus lae
76	7	1.9	491	1	AMPG_ECOLI	P36670 escherichia
77	7	1.9	495	1	P033_MOUSE	P31361 mus musculus
78	7	1.9	497	1	P033_RAT	Q63262 rattus norv
79	7	1.9	500	1	P033_HUMAN	P20264 homo sapien
80	7	1.9	503	1	ATPA_ODOSI	Q00820 odontella s
81	7	1.9	510	1	ATPA_RICCN	Q92986 rickettsia
82	7	1.9	512	1	ATPA_RICPR	O50288 rickettsia
83	7	1.9	527	1	YF00_MYCPN	P75287 mycoplasma
84	7	1.9	528	1	ATPA_STRI1	P50001 streptomyce
85	7	1.9	532	1	ZIC2_HUMAN	O95409 homo sapien
86	7	1.9	536	1	TCPD_FUGRU	P53451 fugu rubrip
87	7	1.9	539	1	TCPD_HUMAN	P50991 homo sapien
88	7	1.9	539	1	TCPD_MOUSE	P80315 mus musculus
89	7	1.9	540	1	TCPD_CABEL	P47208 caenorhabdi
90	7	1.9	549	1	DSK DROME	P23023 drosophila
91	7	1.9	556	1	SR54_CANAL	O42816 candida alb
92	7	1.9	557	1	PKSJ_BACSU	P40806 bacillus su
93	7	1.9	572	1	ACSA_BACSU	P39062 bacillus su
94	7	1.9	576	1	P80C_HUMAN	P38432 homo sapien
95	7	1.9	590	1	K2C5_HUMAN	P13647 homo sapien
96	7	1.9	612	1	ADP1_CANAL	P46589 candida alb
97	7	1.9	622	1	KICI_HUMAN	P35527 homo sapien
98	7	1.9	624	1	STS_MOUSE	P50427 mus musculus
99	7	1.9	627	1	TES1_MOUSE	O70146 mus musculus
100	7	1.9	628	1	TES1_RAT	Q63572 rattus norv

```

RH23_SCHPO STANDARD; PRT; 368 AA.
ID RH23_SCHPO STANDARD; PRT; 368 AA.
AC O74803;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE UV excision repair protein rhp23 (RAD23 homolog).
GN RHP23 OR SPBC2D10.12.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
RC STRAIN=SP223;
RX MEDLINE=21648918; PubMed=11788722;
RA Elder R.T., Song X.-Q., Chen M., Hopkins K.M., Lieberman H.B.,
RA Zhao Y.;
RT "Involvement of rhp23, a Schizosaccharomyces pombe homolog of the
RT human HHR23A and Saccharomyces cerevisiae RAD23 nucleotide excision
RT repair genes, in cell cycle control and protein ubiquitination.";
RL Nucleic Acids Res. 30:581-591(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Spurois J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churchman C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Gobie A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=20119127; PubMed=10652237;
RA Lombaerts M., Goeloe J.I., den Dulk H., Brandema J.A., Brouwer J.;
RT "Identification and characterization of the rhp23(+) DNA repair gene
RT in Schizosaccharomyces pombe.";
RL Biochem. Biophys. Res. Commun. 268:210-215(2000).
CC -1- FUNCTION: Involved in postreplication repair of UV-damaged DNA.
CC strand on replication functions in gap-filling of a daughter
CC strand on replication of damaged DNA.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Contains 1 ubiquitin-like domain.
CC -1- SIMILARITY: Contains 2 UBA domains.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/

```

```

or send an email to license@isb-sib.ch).
CC EMBL; AF174293; AAD51975.1; -.
CC EMBL; AL031788; CAA21170.1; -.
DR PIR; T40115; T40115.
DR HSP; P54725; IDV0.
DR GeneDB_Sfombe; SPBC2D10.12; -.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006636; STII.
DR InterPro; IPR000449; UBA domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00727; STII; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBQ; 1.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS50030; UBA; 2.
DR PROSITE; PS50053; UBIQUITIN 2; 1.
KW DNA damage; DNA repair; Nuclear protein; Repeat.
FT DOMAIN 1 77 UBIQUITIN-LIKE.
FT DOMAIN 135 185 UBA 1.
FT DOMAIN 320 360 UBA 2.
FT DOMAIN 119 122 POLY-ALA.
FT DOMAIN 205 208 POLY-GLN.
SQ SEQUENCE 368 AA; 40135 MW; 5CE75EB7E190EFD4 CRC64;

Query Match 2.4%; Score 9; DB 1; Length 368;
Best Local Similarity 100.0%; Pred. No. 0.42; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

QY 352 NEELAANYL 360
DB 349 NEELAANYL 357
|||||||
|||||||

RESULT 2
SHX2_HUMAN STANDARD; PRT; 331 AA.
ID SHX2_HUMAN STANDARD; PRT; 331 AA.
AC O60902; O60465; O60467; O60903;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Short stature homeobox protein 2 (Paired-related homeobox protein
DE SHOT) (Homeobox protein Ogl2X).
GN SHOX2 OR SHOT OR OG12X.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Muscle;
RX MEDLINE=98151525; PubMed=9482898;
RA Blaschke R.J., Monaghan A.P., Schiller S., Schechinger B., Rao E.,
RA Padilla-Nash H., Ried T., Rappold G.A.;
RT "SHOT, a SHOX-related homeobox gene, is implicated in craniofacial,
RT brain, heart, and limb development.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:2406-2411(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uadiri T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

```

FT	DNA_BIND	140	199	HOMEBOX.
FT	DOMAIN	313	326	OAR.
FT	DOMAIN	60	86	POLY-GLY.
FT	VARSPLOC	235	246	Missing (in isoform 2).
FT				/FTId=VSP_002288.
FT	CONFLICT	115	115	E -> EGRRPTTAAEQVATLLPGENAPRL (IN REF. 2).
FT	CONFLICT	125	125	E -> D (IN REF. 1; CAA05341).
FT	CONFLICT	244	244	P -> S (IN REF. 2).
FT	CONFLICT	312	312	D -> N (IN REF. 2 AND 3).
FT	CONFLICT	325	325	H -> L (IN REF. 3; AACG39663).
SEQ	SEQUENCE	331 AA;	34964 MW;	55431B073B3B2250 CRG64;

Query Match 2.2%; Score 8; DB 1; Length 331;
 Best Local Similarity 100.0%; Pred. No. 4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 234 AGGAGGG 241
 |||||
 Db 83 AGGAGGG 90

RESULT 3

SHX2_MOUSE	STANDARD;	PRT;	331 AA.
AC	P70390; P70369;		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	Short stature homeobox protein 2 (Homeobox protein Ogl2X) (OG-12)		
DE	(Paired family homeodomain protein Prx3).		
GN	SHOX2 OR OGL2X OR PRX3.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxId=10090;		
RI	[1]		
RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).		
RC	TISSUE=Embryo;		
RX	MEDLINE=98058757; PubMed=9371788;		
RA	van Schaick H.S.A., Smidt M.P., Rovescalli A.C., Luijten M.,		
RA	van der Kleij A.A.M., Asoh S., Kozak C.A., Nirenberg M.W.,		
RA	Burbach J.P.H.;		
RT	"Homeobox gene Prx3 expression in rodent brain and extraneural		
RT	tissues."		
RL	Proc. Natl. Acad. Sci. U.S.A. 94:12993-12998(1997).		
RN	[2]		
RP	SEQUENCE OF 116-331 AND 235-331 FROM N.A. (ISOFORMS 1 AND 2).		
RC	STRAIN=BALB/C; TISSUE=Embryo, and Liver;		
RX	MEDLINE=97008065; PubMed=8855241;		
RA	Rovescalli A.C., Asoh S., Nirenberg M.W.;		
RT	"Cloning and characterization of four murine homeobox genes.;"		
RL	Proc. Natl. Acad. Sci. U.S.A. 93:10691-10696(1996).		
RN	[3]		
RP	DEVELOPMENTAL EXPRESSION.		
RX	MEDLINE=98133920; PubMed=9466998;		
RA	Semina E.V., Reiter R.S., Murray J.C.;		
RT	"A new human homeobox gene OGL2X is a member of the most conserved		
RT	homeobox gene family and is expressed during heart development in		
RT	mouse."		
RL	Hum. Mol. Genet. 7:415-422(1998).		
CC	-1- FUNCTION: May be a growth regulator and have a role in specifying		
CC	neural systems involved in processing somatosensory information, as		
CC	as well as in face and body structure formation. May also have a		
CC	role in heart development.		
CC	-1- SUBCELLULAR LOCATION: Nuclear.		
CC	-1- ALTERNATIVE PRODUCTS:		
CC	Event=Alternative splicing; Named isoforms=2;		
CC	Name=1; Synonyms=OGL2A, PRX3A;		
CC	Isoid=P70390-1; Sequence=Displayed;		
CC	Name=2; Synonyms=OGL2B, PRX3B;		
CC	Isoid=P70390-2; Sequence=VSP_002289;		
CC	-1- TISSUE SPECIFICITY: Highly expressed in striated muscle followed		

```

NCBI_TaxID=3702;
[1]
SEQUENCE FROM N.A.
STRAIN=cv. Columbia;
MEDLINE=92003708; PubMed=1912511;
Quigley F., Villiot M.L., Mache R.;
"Nucleotide sequence and expression of a novel glycine-rich protein
gene from Arabidopsis thaliana.";
Plant Mol. Biol. 17:949-952(1991).
[2]
SEQUENCE FROM N.A.
STRAIN=cv. Columbia;
MEDLINE=20277480; PubMed=10819329;
Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;
"Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
features of the regions of 4,504,864 bp covered by sixty P1 and TAC
clones.";
DNA Res. 7:131-135(2000).
-1- FUNCTION: RESPONSIBLE FOR PLASTICITY OF THE CELL WALL (POTENTIAL).
-1- SUBCELLULAR LOCATION: Cell wall (Potential).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR ENBL; X58338; CAA41249.1; ALT_INIT.
ENBL; AB026636; BAA94983.1; -.
KW Cell wall; Structural protein; Repeat; Signal.
FT SIGNAL 1 23 POTENTIAL.
FT TAGN 24 349 GLYCINE-RICH CELL WALL STRUCTURAL
FT FT FT PROTEIN.
FT FT FT
FT DOMAIN 21 338 GLY-RICH.
FT SEQUENCE 349 AA; 25029 MW; 386ADC6817621D46 CRC64;
SQ
Query Match 2.2%; Score 8; DB 1; Length 349;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 234 AGGAGGG 241
|||||||
Db 63 AGGAGGG 70
RESULT 5
GRP1_PETHY STANDARD; PRT; 384 AA.
AC P09789;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Glycine-rich cell wall structural protein 1 precursor.
GN GRP-1.
OS Petunia hybrida (Petunia).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Petunia.
OC NCBI_TaxID=4102;
RN [1]
RP SEQUENCE FROM N.A.
RA Condit C.M., Meagher R.B.;
RT "A gene encoding a novel glycine-rich structural protein of petunia.";
RL Nature 323:178-181(1986).
CC -1- FUNCTION: RESPONSIBLE FOR PLASTICITY OF THE CELL WALL (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Cell wall (Potential).
CC -1- MISCELLANEOUS: THIS PROTEIN CONTAINS 67% GLYCINE RESIDUES.
CC -1- MISCELLANEOUS: 90% OF THE MATURE PROTEIN RESIDUES ARE CAPABLE OF
FORMING A BETA-PLATED SHEET COMPOSED OF 8 ANTI-PARALLEL STRANDS.
CC -1- MISCELLANEOUS: THE GLYCINE-RICH REGION IS COMPRISED OF TWO RELATED
FAMILIES OF REPEATS, F1 AND F2, EACH REPEAT CONTAINING ABOUT 40

```

```

CC AA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X04335; CAA27866.1; -.
CC DR HSP: A26099; A26099.
CC DR PIR: A26099; A26099.
CC KW Cell wall; Structural protein; Repeat; Signal.
CC FT SIGNAL 1 27
CC FT CHAIN 28 384 GLYCINE-RICH CELL WALL STRUCTURAL
CC FT DOMAIN 41 384 GLY-RICH.
CC FT SEQUENCE 384 AA; 28777 MW; C8541C549417D18C CRC64;
CC
CC Query Match 2.2%; Score 8; DB 1; Length 384;
CC Best Local Similarity 100.0%; Pred. No. 4.6;
CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 234 AGGAGGG 241
CC DB 75 AGGAGGG 82
CC
CC RESULT 6
CC FXD1 MOUSE
CC ID FXD1 MOUSE STANDARD; PRT; 456 AA.
CC AC Q61345;
CC DT 15-JUL-1998 (Rel. 36, Created)
CC DT 15-JUL-1998 (Rel. 36, Last sequence update)
CC DT 30-MAY-2000 (Rel. 39, Last annotation update)
CC DE Forkhead box protein D1 (Forkhead-related protein FXH1L8) (Forkhead-
CC related transcription factor 4) (FREAC-4).
CC GN FOXD1 OR FXH1L8 OR FREAC4 OR FHBF2.
CC OS Mus musculus (Mouse).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC OX NCBI_TaxID=10090;
CC RN [1]_TaxID=10090;
CC RP SEQUENCE FROM N.A.
CC RC STRAIN-BALB/C; TISSUE=Embryo;
CC RX MEDLINE=95114592; PubMed=7815060;
CC RA Hatini V., Tao W., Lai E.;
CC RT "Expression of winged helix genes, BF-1 and BF-2, define adjacent
CC domains within the developing forebrain and retina.";
CC RL J. Neurobiol. 25:1293-1309(1994).
CC CC - FUNCTION: MAY FUNCTION IN SPECIFYING POSITIONAL IDENTITY IN THE
CC DEVELOPING RETINA AS WELL AS THE SUBDIVISION OF THE FOREBRAIN
CC NEUROEPITHELIUM.
CC CC - SUBCELLULAR LOCATION: Nuclear (Potential).
CC CC - TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE CNS AND
CC TEMPORAL HALF OF THE RETINA. ALSO EXPRESSED IN THE CONDENSED HEAD
CC MESENCHYME, METANEPHRIC BLASTOMA OF THE DEVELOPING KIDNEY, CORTEX
CC OF THE ADRENAL GLAND, CONDENSED MESENCHYME AT THE BASE OF THE
CC FOLLICLES OF VIBRASSAE, AND CARTILAGE PERICHONDRUM OF THE
CC DEVELOPING VERTEBRATE.
CC CC - DEVELOPMENTAL STAGE: AT E9.5 EMBRYOS, EXPRESSED IN A LIMITED
CC REGION OF THE NEUROEPITHELIUM AND ALSO IN THE TEMPORAL HALF OF THE
CC PRIMARY OPTIC CUP AND THE OPTIC STALK. AT E10.5, SEEN IN THE
CC HYPOTHALAMUS, TEMPORAL HALF OF THE OPTIC STALK, AND TEMPORAL
CC HEMIRETINA. AT E12.5 AND E13.5 A HIGH EXPRESSION IS SEEN IN
CC REGIONS OF CONDENSED MESENCHYME OF THE HEAD, AND AS
CC NEUROEPITHELIAL CELLS BEGIN TO DIFFERENTIATE AND MIGRATE OUTWARD
CC FROM THE VENTRICULAR ZONE, EXPRESSION DECLINES MARKEDLY. BY E16.5
CC LEVELS ARE DIMINISHED AND RESTRICTED TO UNFUSED POCKETS ALONG THE
CC EXHAUSTED VENTRICULAR ZONE.
CC CC - SIMILARITY: Contains 1 fork-head domain.
CC CC - CAUTION: WAS ORIGINALLY (REF.1) ASSIGNED TO BE BF-2 (FOXG1A).
CC -----

```

```

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L38607; AAC42042.1; -.
CC DR HSP: O63245; 2HFH.
CC DR TRANSFAC: T02293; -.
CC DR MGD; MGI:1347463; Foxd1.
CC DR InterPro; IPR001766; TP_Fork_head.
CC DR Pfam; PF00250; Fork_head; 1.
CC DR PRINTS; PR00053; FORKHEAD.
CC DR ProDom; PD000425; TP_Fork_head; 1.
CC DR SMART; SM00339; FH; 1.
CC DR PROSITE; PS00657; FORK_HEAD_1; 1.
CC DR PROSITE; PS00658; FORK_HEAD_2; 1.
CC DR PROSITE; PS50039; FORK_HEAD_3; 1.
CC KW Transcription regulation; DNA-binding; Nuclear protein;
CC Developmental protein.
CC FT DOMAIN 29 32 POLY-GLU.
CC FT DOMAIN 33 36 POLY-ASP.
CC FT DOMAIN 38 46 POLY-GLY.
CC FT DOMAIN 55 58 POLY-ARG.
CC FT DOMAIN 73 76 POLY-ASP.
CC FT DNA_BIND 130 224 FORK-HEAD.
CC FT DOMAIN 261 267 POLY-PRO.
CC FT DOMAIN 293 302 POLY-ALA.
CC FT DOMAIN 308 319 POLY-PRO.
CC FT DOMAIN 395 400 POLY-GLY.
CC FT DOMAIN 420 425 POLY-ALA.
CC SQ SEQUENCE 456 AA; 45429 MW; 43D93F89BBDPCCC7 CRC64;
CC
CC Query Match 2.2%; Score 8; DB 1; Length 456;
CC Best Local Similarity 100.0%; Pred. No. 5.5;
CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 234 AGGAGGG 241
CC DB 108 AGGAGGG 115
CC
CC RESULT 7
CC GRP2 PHAVU STANDARD; PRT; 465 AA.
CC ID GRP2 PHAVU STANDARD; PRT; 465 AA.
CC AC P10496;
CC DT 01-JUL-1989 (Rel. 11, Created)
CC DT 01-JUL-1989 (Rel. 11, Last sequence update)
CC DT 01-OCT-1994 (Rel. 30, Last annotation update)
CC DE Glycine-rich cell wall structural protein 1.8 precursor (GRP 1.8).
CC OS Phaseolus vulgaris (Kidney bean) (French bean).
CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
CC OC eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
CC OX NCBI_TaxID=3885;
CC RN [1]_TaxID=3885;
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=cv. Tendergreen;
CC RX MEDLINE=89091109; PubMed=3208742;
CC RA Keller B., Sauer N., Lamb C.J.;
CC RT "Glycine-rich cell wall proteins in bean: gene structure and
CC association of the protein with the vascular system.";
CC RL EMBO J. 7:3625-3633(1988).
CC CC - FUNCTION: RESPONSIBLE FOR PLASTICITY OF THE CELL WALL (POTENTIAL).
CC CC - SUBCELLULAR LOCATION: Cell wall (Potential).
CC CC - DOMAIN: THE REPEATED DOMAINS OF THE PROTEIN FORM A BETA-PLATED
CC SHEET CONFIGURATION.
CC CC - SIMILARITY: THE N-TERMINAL SIGNAL SEQUENCE OF ABOUT 30 AA SHOWS
CC ABOUT 60% HOMOLOGY TO THAT OF THE GLYCINE-RICH CELL WALL
CC STRUCTURAL PROTEIN GRP 1.0.
CC -----

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL: X13596; CAA31932.1; --
CC PIR: S01820; S01820.
CC Cell wall; Structural protein; Repeat; Signal.
KW SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 465 GLYCINE-RICH CELL WALL STRUCTURAL
FT PROTEIN 1.8.
FT DOMAIN 33 465 GLY-RICH.
FT DOMAIN 205 359 8 X 22 AA TANDEM REPEATS.
FT SEQUENCE 465 AA; 36683 MW; B5C4A9B983B43607 CRC64;

Query Match 2.2%; Score 8; DB 1; Length 465;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGGG 241
| | | | |
DB 191 AGGAGGG 198

RESULT 8

FXD2 HUMAN STANDARD; PRT; 497 AA.
AC O60548;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Forkhead box protein D2 (forkhead-related protein FKHL17) (Forkhead-
DE related transcription factor 9) (FREAC-9).
GN FOXD2 OR FKHL17 OR FREAC9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98066765; PubMed=9403061;
RA Ernstson S., Betz R., Lagercrantz S., Larsson C., Erickson S.,
RA Cederberg A., Carlsson P., Enerbaeck S.;
RT "Cloning and characterization of freac-9 (FKHL17), a novel kidney-
RT expressed human forkhead gene that maps to chromosome 1p32-p34.";
RL Genomics 46:78-85(1997).
RN [2]
RP REVISIONS
RA Enerbaeck S.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PROBABLE TRANSCRIPTION FACTOR.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: KIDNEY-SPECIFIC.
CC -1- SIMILARITY: Contains 1 fork-head domain.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL: AF042832; AAC15421.1; --
CC HSP: Q63245; 2HF5.
CC TRANSFAC: T02485;
CC Genew: HGNC:3803; FOXD2.
CC MM: 602211; --
CC GO: GO:0003700; F:transcription factor activity; TAS.
CC InterPro: IPR001766; TF_Fork_head.

DR Pfam; PF00250; Fork head; 1.
DR PRINTS; PR00053; FORKHEAD.
DR ProDom; PD000425; TF_Fork_head; 1.
DR SMART; SM00339; FH; 1.
DR PROSITE; PS00657; FORK_HEAD_1; 1.
DR PROSITE; PS00658; FORK_HEAD_2; 1.
DR PROSITE; PS00339; FORK_HEAD_3; 1.
KW DNA-binding; Nuclear protein; Transcription regulation.

FT DOMAIN 90 94 POLY-ALA.
FT DOMAIN 101 104 POLY-ALA.
FT DNA_BIND 126 217 FORK-HEAD.
FT DOMAIN 247 250 POLY-ALA.
FT DOMAIN 296 306 POLY-ALA.
FT DOMAIN 398 409 POLY-GLY.
FT DOMAIN 421 426 POLY-GLY.
FT DOMAIN 442 445 POLY-ALA.
SQ SEQUENCE 497 AA; 49007 MW; EAAF498D216BE019 CRC64;

Query Match 2.2%; Score 8; DB 1; Length 497;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGGG 241
| | | | |
DB 397 AGGAGGG 404

RESULT 9

IRBP_MOUSE STANDARD; PRT; 610 AA.
AC P49194; Q9ROH8;
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Interphotoreceptor retinoid-binding protein precursor (IRBP)
DE (interstitial retinoid-binding protein) (Fragment).
GN RBP3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Si J.S., Nickerson J.M.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 42-454 FROM N.A., AND TISSUE SPECIFICITY.
RC STRAIN=BALB/c;
RX MEDLINE=94115692; PubMed=1342928;
RA Stanhope M.J., Czelusniak J., Si J.-S., Nickerson J., Goodman M.;
RT "A molecular perspective on mammalian evolution from the gene encoding
RT interphotoreceptor retinoid binding protein, with convincing evidence
RT for bat monophyly.";
RL Mol. Phylogenet. Evol. 1:148-160(1992).
RN [3]
RP SEQUENCE OF 71-454 FROM N.A.
RC STRAIN=MOA;
RA Suzuki H., Serizawa K.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: IRBP SHUTTLES 11-CIS AND ALL TRANS RETINOLIDS BETWEEN
CC THE RETINOL ISOMERASE IN THE PIGMENT EPITHELIUM AND THE VISUAL
CC PIGMENTS IN THE PHOTORECEPTOR CELLS OF THE RETINA.
CC -1- SUBCELLULAR LOCATION: INTERPHOTORECEPTOR MATRIX THAT PERMEATES
CC THE SPACE BETWEEN THE RETINA AND THE CONTIGUOUS LAYER OF PIGMENT
CC EPITHELIUM CELLS.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE PHOTORENSITIVE TISSUES;
CC RETINA AND PINEAL GLAND.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF126968; AAA39331.2; --
 DR EMBL; AB033711; BAA85872.1; --
 DR MGD; MGI:97878; Rbp3.
 DR InterPro; IPR003581; TSPc.
 DR Pfam; PF02692; IRBP; 2.
 DR SMART; SM00245; TSPC; 2.
 KW Vitamin A; Transport; Repeat; Signal; Glycoprotein.
 FT SIGNAL 1 17
 FT CHAIN 18 >610
 FT DOMAIN 18 >610
 FT REPEAT 18 320
 FT REPEAT 321 >610
 FT CARBOHYD 107 >610
 FT CARBOHYD 205 205
 FT CARBOHYD 513 513
 FT CONFLICT 161 161
 FT NON TER 610 610
 SQ SEQUENCE 610 AA; 66679 MW; F392FE2D187C6416 CRC64;

Query Match 2.2%; Score 8; DB 1; Length 610;
 Best Local Similarity 100.0%; Pred. No. 7.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 102 LEAPQAP 109
 DB 96 LEAPQAP 103

RESULT 10

ID YC1Q ECOLI STANDARD; PRT; 631 AA.
 AC P45848; P45849; P76031;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein yciQ.
 GN YC1Q OR B1268.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / W3110;
 RA Milkman R., McKane M.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MGI655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97251357; PubMed=9097039;
 RA Alba H., Baba T., Fujita K., Hayaashi K., Inada T., Isono K., Itoh T.,
 RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 RA Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshina T., Saito N.,
 RA Samesi G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
 RA Takeuchi K., Takeuchi Y., Wada C., Yamamoto I., Horiuchi T.;
 RT "A 570-Kb DNA sequence of the Escherichia coli K-12 genome
 corresponding to the 28.0-40.1 min region on the linkage map.";

RL DNA Res. 3:363-377(1996).
 CC -!- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
 CC FRAMESHIFT IN POSITION 198.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U18111; AB59988.1; ALT_FRAME.
 DR EMBL; U18111; AB59988.1; ALT_FRAME.
 DR EMBL; AB000224; AAC74350.1; --
 DR EMBL; D90764; BAA14803.1; ALT_INIT.
 DR EMBL; D90765; BAA14818.1; ALT_INIT.
 DR PIR; G54874; G64874.
 DR EcGene; EG12969; yciQ.
 KW Hypothetical protein; Complete proteome.
 FT DOMAIN 616 630 POLY-GLY.
 FT CONFLICT 84 84 A -> G (IN REF. 1).
 FT CONFLICT 88 88 N -> D (IN REF. 1).
 SQ SEQUENCE 631 AA; 71289 MW; 1154FF72FDC6C16E CRC64;

Query Match 2.2%; Score 8; DB 1; Length 631;
 Best Local Similarity 100.0%; Pred. No. 7.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
 DB 619 AGGGAGGG 626

RESULT 11

ID CADM_HUMAN STANDARD; PRT; 828 AA.
 AC Q9UJ99; O43205;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cadherin-22 precursor.
 GN CDH22 OR C20ORF25.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.P., Bates K.N., Beard L.M., Beare D.M., B.J.,
 RA Beasley O.P., Bird C.P., Blake S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grahame D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Lehmaeslaih M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McLeay K., McMurray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,

RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.; The DNA sequence and comparative analysis of human chromosome 20.";
 RT Nature 414:865-871 (2001).
 RL [2]
 RN SEQUENCE OF 449-828 FROM N.A.
 RP TISSUE=Brain;
 RC Yu W., Sarginson J., Gibbs R.A.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
 CC They preferentially interact with themselves in a homophilic
 CC manner in connecting cells; cadherins may thus contribute to the
 CC sorting of heterogeneous cell types. PB-cadherins may have a role
 CC in the morphological organization of pituitary gland and brain
 CC tissues (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -1- SIMILARITY: Contains 5 cadherin domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC ENBL; AL031687; CAB51587.2; .
 CC EMBL; AF035300; AAB88183.1; .
 CC Genew; HGNC:13251; CDH22.
 CC HSSP; P15116; INCUJ.
 CC InterPro; IPR002126; Cadherin.
 CC InterPro; IPR002233; Cadherin_C_term.
 CC Pfam; PF00028; Cadherin; 5.
 CC Pfam; PF01049; Cadherin_C_term; 1.
 CC PRINTS; P00205; CADHERIN.
 CC SMART; SM00112; CA; 5.
 CC PROSITE; PS00212; CADHERIN 1; 2.
 CC PROSITE; PS02268; CADHERIN 2; 5.
 CC Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
 KW Signal.
 KW SIGNAL 1 34
 KW CHAIN 35 828
 KW DOMAIN 36 624
 KW TRANSNEM 625 645
 KW DOMAIN 646 828
 KW DOMAIN 64 168
 KW DOMAIN 169 277
 KW DOMAIN 278 394
 KW DOMAIN 395 498
 KW DOMAIN 499 616
 KW CARBOHYD 162 162
 KW CARBOHYD 466 466
 KW CARBOHYD 612 612
 KW SEQUENCE 828 AA; 89091 MW; 520F7B1DF624DCA CRC64;
 Query Match 2.2%; Score 8; DB 1; Length 828;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 234 AGGGAGGG 241
 Db 706 AGGGAGGG 713
 RESULT 12
 KDGE DROME
 ID_KDGE DROME STANDARD; PRT; 1454 AA.
 AC Q09103;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Eye-specific diacylglycerol kinase (EC 2.7.1.107) (Retinal
 DE degeneration A protein) (Diglyceride kinase 2) (DGK 2) (DAG kinase 2).

GN RDGA OR DGK2.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=Canton-S; TISSUE=Head;
 RX MEDLINE=9468563; PubMed=8248222;
 RA Masai I., Okazaki A., Hosoya T., Hotta Y.;
 RT "Drosophila retinal degeneration A gene encodes an eye-specific
 RT diacylglycerol kinase with cysteine-rich zinc-finger motifs and
 RT ankyrin repeats.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:11157-11161 (1993).
 CC -1- FUNCTION: REQUIRED FOR THE MAINTENANCE OF THE PHOTORECEPTOR. ITS
 CC ABSENCE LEADS TO RHADOMERE DEGENERATION DUE TO DEFECTIVE
 CC PHOSPHOLIPID TURNOVER.
 CC -1- CATALYTIC ACTIVITY: ATP + 1,2-diacylglycerol = ADP + 1,2-
 CC diacylglycerol 3-phosphate.
 CC -1- SUBCELLULAR LOCATION: Membrane-associated.
 CC -1- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN ADULT EYE.
 CC -1- DISEASE: MUTANTS OF THIS GENE ARE CHARACTERIZED BY HAVING
 CC PHOTORECEPTOR CELLS THAT DEVELOP NORMALLY BUT DEGENERATE RAPIDLY
 CC AFTER ECLOSION.
 CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC DIACYLGLYCEROL KINASE
 CC FAMILY.
 CC -1- SIMILARITY: Contains 2 zinc-dependent phorbol-ester and DAG
 CC binding domains.
 CC -1- SIMILARITY: Contains 4 ANK repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC ENBL; D17315; BAA04135.1; .
 CC PIR; T13709; T13709.
 CC FlyBase; FBgn0003217; rdga.
 CC GO; GO:0016059; P:deactivation of rhodopsin mediated signaling; IGI.
 CC GO; GO:0016056; P:rhodopsin mediated signaling; IMP.
 CC InterPro; IPR002110; ANK.
 CC InterPro; IPR002219; DAG_PE-bind.
 CC InterPro; IPR000756; DAGKa.
 CC InterPro; IPR001206; DAGKc.
 CC Pfam; PF00023; ank; 4.
 CC Pfam; PF00130; DAG_PE-bind; 1.
 CC Pfam; PF00609; DAGKa; 1.
 CC Pfam; PF00781; DAGKc; 1.
 CC ProDom; PD002939; DAGKa; 1.
 CC ProDom; PD005043; DAGKc; 1.
 CC SMART; SM00248; ANK; 2.
 CC SMART; SM00109; C1; 2.
 CC SMART; SM00045; DAGKa; 1.
 CC SMART; SM00046; DAGKc; 1.
 CC PROSITE; PS50088; ANK_REPEAT; 2.
 CC PROSITE; PS00297; ANK_REPEAT_REGION; 1.
 CC PROSITE; PS00479; DAG_PE_BIND_DOM_1; FALSE NEG.
 CC PROSITE; PS50081; DAG_PE_BIND_DOM_2; FALSE NEG.
 KW transferase; Kinase; ANK repeat; Repeat; Vision;
 KW Phorbol-ester binding.
 KW DOMAIN 592 642
 KW DOMAIN 662 719
 KW DOMAIN 807 935
 KW DOMAIN 961 1115
 KW REPEAT 1317 1346
 KW REPEAT 1350 1379
 KW REPEAT 1386 1415
 KW REPEAT 1419 1448
 KW DOMAIN 2 6
 KW PHORBOL-ESTER AND DAG BINDING 1.
 KW PHORBOL-ESTER AND DAG BINDING 2.
 KW CATALYTIC-A (POTENTIAL).
 KW CATALYTIC-B (POTENTIAL).
 KW ANK 1.
 KW ANK 2.
 KW ANK 3.
 KW ANK 4.
 KW POLY-GLN.

```

FT DOMAIN 24 39 THR-RICH.
FT DOMAIN 110 115 POLY-SER.
FT DOMAIN 227 231 POLY-GLU.
FT DOMAIN 758 775 GLY-RICH.
FT MUTAGEN 869 869 G->D: IN MUTANT RDGA2.
SQ SEQUENCE 1454 AA; 159675 MW; C84C81095PEA16AA CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGGG 241
Db 764 AGGAGGG 771

RESULT 13
NEUG_BOVIN
ID NEUG_BOVIN STANDARD; PRT; 76 AA.
AC F35722;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurogranin (NG) (P17) (B-50 immunoreactive C-kinase substrate)
DE (BICKS) (Fragment).
GN NRGN.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=91093129; PubMed=1824695;
RA Baudier J., Deloume J.C., van Dorsselaer A., Black D.,
RT "Purification and characterization of a brain-specific protein kinase
RT C substrate, neurogranin (p17). Identification of a consensus amino
RT acid sequence between neurogranin and neuromodulin (GAP43) that
RT corresponds to the protein kinase C phosphorylation site and the
RT calmodulin-binding domain."
RL J. Biol. Chem. 266:229-237(1991).
CC -!- FUNCTION: ACTS AS A "THIRD MESSENGER" SUBSTRATE OF PROTEIN KINASE
CC C-MEDIATED MOLECULAR CASCADES DURING SYNAPTIC DEVELOPMENT AND
CC REMODELING. BINDS TO CALMODULIN IN THE ABSENCE OF CALCIUM.
CC -!- TISSUE SPECIFICITY: IS HIGHLY ENRICHED IN BRAIN ACCUMULATES
CC POSTSYNAPTICALLY IN DENDRITIC SPINES OF NEOSTRIATAL NEURONS.
CC -!- MASS SPECTROMETRY: MW=7837.1; MW ERR=0.5; METHOD=Electrospray.
CC -!- SIMILARITY: BELONGS TO THE NEUROGRANIN FAMILY.
CC PIR; A39034; A39034.
DR InterPro; IPR000048; IQ_region.
DR Pfam; PF00612; IQ; 1.
DR SMART; SM00015; IQ; 1.
DR PROSITE; PS50096; IQ; 1.
KW Calmodulin-binding; Phosphorylation; Neurone.
FT DOMAIN 24 47 IQ.
FT DOMAIN 48 76 COLLAGEN-LIKE.
FT MOD_RES 34 34 PHOSPHORYLATION (BY PKC).
FT NON_TER 76 76
SQ SEQUENCE 76 AA; 7248 MW; DA898C93FD6165CF CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 GGAGGGP 242
Db 67 GGAGGGP 73

RESULT 14
NEUG_BOVIN
ID NEUG_BOVIN STANDARD; PRT; 76 AA.
AC Q92686;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Neurogranin (NG) (RC3).
GN NRGN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97288522; PubMed=9143500;
RA Martinez de Arrieta C., Perez Jurado L., Bernal J., Coloma A.;

```

```

NEUG_CAPHI
ID NEUG_CAPHI STANDARD; PRT; 78 AA.
AC P54877;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurogranin (NG) (Protein kinase C substrate 7.5 kDa protein) (RC3).
GN NRGN.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95287801; PubMed=7539519;
RA Piosik P.A., van Groenigen M., Ponne N.J., Bolhuis P.A., Baas F.;
RT "RC3/neurogranin structure and expression in the caprine brain in
RT relation to congenital hypothyroidism."
RL Brain Res. Mol. Brain Res. 29:119-130(1995).
CC -!- FUNCTION: ACTS AS A "THIRD MESSENGER" SUBSTRATE OF PROTEIN KINASE
CC C-MEDIATED MOLECULAR CASCADES DURING SYNAPTIC DEVELOPMENT AND
CC REMODELING. BINDS TO CALMODULIN IN THE ABSENCE OF CALCIUM (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE NEUROGRANIN FAMILY.
CC -!- SIMILARITY: Contains 1 IQ domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; S78295; AAB34481.1; -.
DR InterPro; IPR000048; IQ_region.
DR Pfam; PF00612; IQ; 1.
DR SMART; SM00015; IQ; 1.
DR PROSITE; PS50096; IQ; 1.
KW Calmodulin-binding; Phosphorylation; Neurone.
FT DOMAIN 26 47 IQ.
FT DOMAIN 48 78 COLLAGEN-LIKE.
FT MOD_RES 36 36 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
SQ SEQUENCE 78 AA; 7549 MW; 8E47CDA39F085794 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 GGAGGGP 242
Db 69 GGAGGGP 75

RESULT 15
NEUG_HUMAN
ID NEUG_HUMAN STANDARD; PRT; 78 AA.
AC Q92686;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Neurogranin (NG) (RC3).
GN NRGN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97288522; PubMed=9143500;
RA Martinez de Arrieta C., Perez Jurado L., Bernal J., Coloma A.;

```

RT "Structure, organization, and chromosomal mapping of the human
 RT neurogranin gene (NRGN).";
 RL Genomics 41:243-249 (1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97041074; PubMed=8929222;
 RA Mertsalov I.B., Gundelfinger E.D., Tsetlin V.I.;
 RT "Cloning cDNA for human neurogranin.";
 RL Bioorg. Khim. 22:366-369 (1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Chang J.W., Schumacher E., Coulter P.M. II, Vinters H.V., Watson J.B.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98100924; PubMed=9499372;
 RA Mertsalov I.B., Stumm M., Wieacker P., Tom Dieck S.,
 RA Gundelfinger E.D., Tsetlin V.I.;
 RT "Structure and chromosomal localization of human neurogranin gene.";
 RL Bioorg. Khim. 23:961-968 (1997).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Mertsalov I.B., Tsetlin V.I.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Uterus;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.A., Loquellano N.J., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -!- FUNCTION: ACTS AS A "THIRD MESSENGER" SUBSTRATE OF PROTEIN KINASE
 CC C-MEDIATED MOLECULAR CASCADES DURING SYNAPTIC DEVELOPMENT AND
 CC REMODELING. BINDS TO CALMODULIN IN THE ABSENCE OF CALCIUM (BY
 CC SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE NEUROGRANIN FAMILY.
 CC -!- SIMILARITY: Contains 1 IQ domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X99076; CAA67533.1; -;
 DR EMBL; X99075; CAA67533.1; JOINED.
 DR EMBL; Y09689; CAA70867.1; -;
 DR EMBL; U89165; AAB49458.1; -;
 DR EMBL; Y15058; CAA75343.1; -;
 DR EMBL; Y15059; CAA75343.1; JOINED.
 DR EMBL; AJ317956; CAC37631.1; -;
 DR EMBL; BC002835; AA02835.1; -;
 DR Genew; HGNC:8000; NRGN.

DR MIM; 602350; -;
 DR GO; 0005516; F:calmodulin binding activity; TAS.
 DR GO; 0007399; P:neurogenesis; TAS.
 DR GO; 0007165; P:signal transduction; TAS.
 DR InterPro; IPR000048; IQ_region.
 DR Pfam; PF00612; IQ; 1.
 DR SMART; SM00015; IQ; 1.
 DR PROSITE; P85096; IQ; 1.
 KW Calmodulin-binding; Phosphorylation; Neurone.
 FT DOMAIN 26 47 IQ
 FT DOMAIN 48 78 COLLAGEN-LIKE
 FT MOD_RES 36 36 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
 SQ SEQUENCE 78 AA; 7618 MW; 0B6ED790A31D5785 CRC64;
 Query Match 1.9%; Score 7; DB 1; Length 78;
 Best Local Similarity 100.0%; Pred. No. 9.7;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 236 GGAGGGP 242
 Db 69 GGAGGGP 75

RESULT 16

NEUG RAT
 ID NEUG RAT STANDARD; PRT; 78 AA.
 AC 004940;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurogranin (NG) (Protein kinase C substrate 7.5 kDa protein) (RC3).
 GN NRGN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 RX MEDLINE=91039409; PubMed=2231781;
 RA Watson J.B., Battenberg E.F., Wong K.K., Bloom P.E.,
 RA Sutcliffe J.G.;
 RT "Subtractive cDNA cloning of RC3, a rodent cortex-enriched mRNA
 RT encoding a novel 78 residue protein.";
 RL J. Neurosci. Res. 26:397-408 (1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Testis;
 RX MEDLINE=95247744; PubMed=7730337;
 RA Sato T., Xiao D.M., Li H., Huang F.L., Huang K.P.;
 RT "Structure and regulation of the gene encoding the neuron-specific
 RT protein kinase C substrate neurogranin (RC3 protein).";
 RL J. Biol. Chem. 270:10314-10322 (1995).
 RN [3]
 RP SEQUENCE, AND CHARACTERIZATION.
 RC TISSUE=Brain;
 RX MEDLINE=93384318; PubMed=8080473;
 RA Huang K.-P., Huang F.L., Chen H.-C.;
 RT "Characterization of a 7.5-kDa protein kinase C substrate (RC3
 RT protein, neurogranin) from rat brain.";
 RL Arch. Biochem. Biophys. 305:570-580 (1993).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=92409560; PubMed=1528865;
 RA Watson J.B., Sutcliffe J.G., Fisher R.S.;
 RT "Localization of the protein kinase C
 RT phosphorylation/calmodulin-binding substrate RC3 in dendritic spines
 RT of neostriatal neurons.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:8581-8585 (1992).
 CC -!- FUNCTION: ACTS AS A "THIRD MESSENGER" SUBSTRATE OF PROTEIN KINASE
 CC C-MEDIATED MOLECULAR CASCADES DURING SYNAPTIC DEVELOPMENT AND
 CC REMODELING. BINDS TO CALMODULIN IN THE ABSENCE OF CALCIUM.
 CC -!- TISSUE SPECIFICITY: IS HIGHLY ENRICHED IN BRAIN. ACCUMULATES

CC POSTSYNAPTICALLY IN DENDRITIC SPINES OF NEOSTRIATAL NEURONS.
 CC -1- PTM: PHOSPHORYLATION IS ACTIVATED BY CALCIUM, PHOSPHOLIPID, AND
 CC DIACYLGLYCEROL. PHOSPHORYLATION INHIBITS BINDING TO CALMODULIN
 CC BOTH IN THE PRESENCE AND ABSENCE OF CALCIUM.
 CC -1- SIMILARITY: BELONGS TO THE NEUROGRANIN FAMILY.
 CC -1- SIMILARITY: Contains 1 IQ domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L09119; AAA42023.1; -;
 DR EMBL; U22062; AAA80223.1; -;
 DR PIR; A57288; A57288.
 DR InterPro; IPR000048; IQ_region.
 DR Pfam; PF00612; IQ; 1.
 DR SMART; SM00015; IQ; 1.
 DR PROSITE; PS50096; IQ; 1.
 KW Calmodulin-binding; Phosphorylation; Neurone.
 FT MOD_RES 1 1 BLOCKED.
 FT DOMAIN 26 47 IQ.
 FT DOMAIN 48 78 COLLAGEN-LIKE.
 FT MOD_RES 36 36 PHOSPHORYLATION (BY PKC).
 SQ SEQUENCE 78 AA; 7496 MW; 8B47CDB38E095794 CRC64;
 Query Match 1.9%; Score 7; DB 1; Length 78;
 Best Local Similarity 100.0%; Pred. No. 9.7;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 236 GGAGGGP 242
 DB 69 GGAGGGP 75
 RESULT 17
 Y1B3_CLOAB
 ID Y1B3_CLOAB STANDARD; PRT; 96 AA.
 AC P34159;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein CAC3713.
 GN CAC3713
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 CC Clostridium.
 OX NCBI_TaxID=1488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=93273706; PubMed=8501044;
 RA Sauer U., Duerre P.;
 RT "Sequence and molecular characterization of a DNA region encoding a
 RT small heat shock protein of Clostridium acetobutylicum.";
 RL J. Bacteriol. 175:3394-3400(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=21359325; PubMed=11466286;
 RA Noelling J., Breton G., Omeichenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson R., Lee H.W., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
 RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 RT "Genome sequence and comparative analysis of the solvent-producing
 RT bacterium Clostridium acetobutylicum.";
 RL J. Bacteriol. 183:4823-4838(2001).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X65276; CAA46375.1; -;
 DR EMBL; AB007866; AAK91633.1; -;
 DR PIR; F97355; F97355.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 96 AA; 11003 MW; 0E427F533D1D79DF CRC64;
 Query Match 1.9%; Score 7; DB 1; Length 96;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 319 ISVTPEE 325
 DB 4 ISVTPEE 10
 RESULT 18
 NLPD_YEREN
 ID NLPD_YEREN STANDARD; PRT; 97 AA.
 AC P47764;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Lipoprotein nlpD (Fragment).
 GN NLPD.
 OS Versinia enterocolitica.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Versinia.
 OX NCBI_TaxID=630;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=W1024 / Serotype O:9;
 RX MEDLINE=95247270; PubMed=7729893;
 RA Iriarte M., Stainier I., Cornelis G.R.;
 RT "The rps gene from Versinia enterocolitica and its influence on
 RT expression of virulence factors.";
 RL Infect. Immun. 63:1840-1847(1995).
 CC -1- SUBCELLULAR LOCATION: Attached to the inner membrane by a lipid
 CC anchor (Potential).
 CC -1- SIMILARITY: BELONGS TO THE E. COLI NLPD / HAEMOPHILUS LPPB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U16152; AAC43390.1; -;
 DR MEROPS; M37.UPW; -;
 DR InterPro; IPR002886; Peptidase M37.
 DR InterPro; IPR000437; Prok_lipoProt.
 DR Pfam; PF01551; Peptidase_M37; 1.
 DR PROSITE; PS00013; PROKAR LIPOPROTEIN; PARTIAL.
 KW Inner membrane; Lipoprotein.
 FT NON_TER 1 1
 SQ SEQUENCE 97 AA; 10667 MW; 9E194F9A13D0838E CRC64;
 Query Match 1.9%; Score 7; DB 1; Length 97;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 81 GSTGTSS 87
 DB 68 GSTGTSS 74

RESULT 19

PER_DROER STANDARD; PRT; 110 AA.

AC Q26288;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Period circadian protein (Fragment).

GN PER.

OS *Drosophila erecta* (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7220;

RN [1]

RP SEQUENCE FROM N.A.; PubMed=1487825;

RX MEDLINE=93140158; PubMed=1487825;

RA Peixoto A.A., Costa R., Wheeler D.A., Hall J.C., Kyriacou C.P.;

RT "Evolution of the threonine-glycine repeat region of the period gene in the melanogaster species subgroup of Drosophila.";

RL J. Mol. Evol. 35:411-419(1992).

CC -!- FUNCTION: ESSENTIAL FOR BIOLOGICAL CLOCK FUNCTIONS. DETERMINES THE PERIOD LENGTH OF CIRCADIAN AND ULTRADIAN RHYTHMS; AN INCREASE IN PER DOSAGE LEADS TO SHORTENED CIRCADIAN RHYTHMS AND A DECREASE LEADS TO LENGTHENED CIRCADIAN RHYTHMS. ESSENTIAL FOR THE CIRCADIAN RHYTHMICITY OF LOCOMOTOR ACTIVITY, ECLOSION BEHAVIOR, AND FOR THE RHYTHMIC COMPONENT OF THE MALE COURTSHIP SONG THAT ORIGINATES IN THE THORACIC NERVOUS SYSTEM. THE BIOLOGICAL CYCLE DEPENDS ON THE RHYTHMIC FORMATION AND NUCLEAR LOCALIZATION OF THE TIM-PER COMPLEX. LIGHT INDUCES THE DEGRADATION OF TIM, WHICH PROMOTES ELIMINATION OF PER. NUCLEAR ACTIVITY OF THE HETERODIMER COORDINATIVELY REGULATES PER AND TIM TRANSCRIPTION THROUGH A NEGATIVE FEEDBACK LOOP. BEHAVES AS A NEGATIVE ELEMENT IN CIRCADIAN TRANSCRIPTIONAL LOOP. DOES NOT APPEAR TO BIND DNA, SUGGESTING INDIRECT TRANSCRIPTIONAL INHIBITION (BY SIMILARITY).

CC -!- SUBUNIT: FORMS HETERODIMER WITH TIMELESS (TIM); THE COMPLEX THEN TRANSLLOCATES INTO THE NUCLEUS (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: NUCLEAR AT SPECIFIC PERIODS OF THE DAY. FIRST ACCUMULATES IN THE PERINUCLEAR REGION ABOUT ONE HOUR BEFORE TRANSLLOCATION INTO THE NUCLEUS. INTERACTION WITH TIM IS REQUIRED FOR NUCLEAR LOCALIZATION (BY SIMILARITY).

CC -!- DOMAIN: THE RUN OF GLY-THR IS IMPLICATED IN THE MAINTENANCE OF CIRCADIAN PERIOD AT DIFFERENT TEMPERATURES. DELETION OF THE REPEAT LEADS TO A SHORTENING OF THE COURTSHIP SONG CYCLE PERIOD, AND THUS COULD BE IMPORTANT FOR DETERMINING FEATURES OF SPECIES-SPECIFIC MATING BEHAVIOR (BY SIMILARITY).

CC -!- PTM: PHOSPHORYLATED WITH A CIRCADIAN RHYTHMICITY, PROBABLY BY THE DOUBLE-TIME PROTEIN (DET). PHOSPHORYLATION COULD BE IMPLICATED IN THE STABILITY OF PER MONOMER AND IN THE FORMATION OF HETERODIMER PER-TIM (BY SIMILARITY).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; S53301; AAB25031.2; .

DR FlyBase; FBgn0012277; Dere\per.

KW Biological rhythms; Repeat; Nuclear protein; Phosphorylation.

FT NON_TER 1

FT DOMAIN 33 64 G-T REPEATS.

FT DOMAIN 85 95 POLY-GLY.

FT NON_TER 110 110

SQ SEQUENCE 110 AA; 9850 MW; 2BAA187EF55CDA95 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 110;

Best Local Similarity 100.0%; Pred. No. 14;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGG 241

Db 87 GGGAGG 93

RESULT 20

HBB_DASAK STANDARD; PRT; 141 AA.

AC P56692;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hemoglobin beta chain.

GN HBB.

OS Dasyatis akajei (Red stingray) (Akai).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;

OC Elasmobranchii; Squalia; Hynostomalea; Pristiostoma; Batoidea;

OC Myliobatiformes; Myliobatoidei; Dasyatidae; Dasyatis.

OX NCBI_TaxID=31902;

RN [1]

RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).

RX TISSUE=Blood;

RX MEDLINE=99322407; PubMed=10393295;

RA Chong K.T., Miyazaki G., Morimoto H., Oda Y., Park S.-Y.;

RT "Structures of the deoxy and CO forms of haemoglobin from Dasyatis akajei, a cartilaginous fish.";

RL Acta Crystallogr. D 55:1291-1300(1999).

CC -!- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE VARIOUS PERIPHERAL TISSUES.

CC -!- SUBUNIT: Heterotetramer of two alpha chains and two beta chains.

CC -!- TISSUE SPECIFICITY: Red blood cells.

CC -!- SIMILARITY: Belongs to the globin family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AB023723; BAA75250.1; .

DR PDB; 1CG5; 01-APR-99.

DR PDB; 1CG8; 01-APR-99.

DR InterPro; IPR000971; Globin.

DR Pfam; PF00042; globin; 1.

DR PROSITE; PS01033; GLOBIN; 1.

KW Heme; Oxygen transport; Transport; Erythrocyte;

KW 3D-structure.

FT INIT_MET 0 0

FT METAL 59 59 IRON (HEME DISTAL LIGAND).

FT METAL 88 88 IRON (HEME PROXIMAL LIGAND).

FT HELIX 5 17

FT HELIX 20 34

FT HELIX 36 39

FT TURN 40 41

FT TURN 43 45

FT TURN 46 47

FT TURN 51 52

FT HELIX 54 72

FT TURN 73 75

FT HELIX 77 80

FT TURN 81 81

FT HELIX 82 91

FT TURN 92 92

FT TURN 95 95

FT HELIX 96 113

FT HELIX 114 116

FT HELIX 119 137

FT TURN 138 139

SQ SEQUENCE 141 AA; 16289 MW; DAED4F578904D27B CRC64;

Query Match 1.9%; Score 7; DB 1; Length 141;

Best Local Similarity 100.0%; Pred. No. 18;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 DKVQAL 169
|||||
Db 61 DKVQAL 67

RESULT 21

TWIST1_HUMAN
ID Q15672; Q92487; Q99804; PRT; 202 AA.
AC Q15672; Q92487; Q99804;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Twist related protein 1 (H-twist).
GN TWIST1 OR TWIST.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC TISSUE=Lung;
RA Wang S.M., Pignolo R.J., Rotenberg M.O., Coljee V.W.,
RA Cristofalo V.J., Sierra F.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Placenta;
RA MEDLINE=97148940; PubMed=8995765;
RA Bourgeois P., Stoetzel C., Bolcato-Bellemin A.-L., Mattei M.-G.,
RA Perrin-Schmitt F.;
RT "The human H-twist gene is located at 7p21 and encodes a B-HLH
RL protein that is 96% similar to its murine M-twist counterpart.";
RL Mamm. Genome 7:915-917(1996).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANTS SCS PRO-119, ILE-135 INS AND PRO-139
RP INS.
RX MEDLINE=97141916; PubMed=8988166;
RA Howard T.D., Paznekas W.A., Green E.D., Chiang L.C., Ma N.,
RA Ortiz de Luna R.I., Delgado C.G., Gonzalez-Ramos M., Kline A.D.,
RA Jabs E.W.;
RT "Mutations in TWIST, a basic helix-loop-helix transcription factor,
RT in Saethre-Chotzen syndrome.";
RL Nat. Genet. 15:36-41(1997).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=97358582; PubMed=9215678;
RA Krebs I., Weis I., Hudler M., Rommens J.M., Roth H., Scherer S.W.,
RA Tsui L.-C., Fuchtbauer E.-M., Grzeschik K.-H., Tsuji K., Kunz J.;
RT "Translocation breakpoint maps 5 kb 3-prime from TWIST in a patient
RT affected with Saethre-Chotzen syndrome.";
RL Hum. Mol. Genet. 6:1079-1086(1997).
RN [5]
RP SEQUENCE FROM N.A.
RA Strong C., Graves T., Duckels G., Ozersky P.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udels T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [7]
RP VARIANTS SCS PRO-131 AND PRO-139 INS.
RX MEDLINE=97141917; PubMed=8988167;
RA el Ghourzi V., le Merrier M., Perrin-Schmitt F., Lajeunie E.,
RA Benit P., Renier D., Bourgeois P., Bolcato-Bellemin A.-L.,
RA Munnich A., Bonaventure J.;
RT "Mutations of the TWIST gene in the Saethre-Chotzen syndrome.";
RL Nat. Genet. 15:42-46(1997).
RN [8]
RP VARIANT BGS VAL-156.
RX MEDLINE=21621049; PubMed=11754069;
RA Seto M.L., Lee S.J., Sze R.W., Cunningham M.L.;
RT "Another TWIST on Baller-Gerold syndrome.";
RL Am. J. Med. Genet. 104:323-330(2001).
CC [1]- FUNCTION: PROBABLE TRANSCRIPTION FACTOR, WHICH SEEMS TO BE
CC INVOLVED IN THE NEGATIVE REGULATION OF CELLULAR DETERMINATION AND
CC IN THE DIFFERENTIATION OF SEVERAL LINEAGES INCLUDING MYOGENESIS,
CC OSTEOGENESIS, AND NEUROGENESIS. INHIBITS MYOGENESIS BY
CC SEQUESTERING E PROTEINS, INHIBITING TRANS-ACTIVATION BY MEF2, AND
CC INHIBITING DNA-BINDING BY MYOD THROUGH PHYSICAL INTERACTION. THIS
CC INTERACTION PROBABLY INVOLVES THE BASIC DOMAINS OF BOTH PROTEINS
CC (BY SIMILARITY).
CC [2]- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
CC BHLH PROTEIN. HOMODIMER (BY SIMILARITY).
CC [3]- SUBCELLULAR LOCATION: Nuclear.
CC [4]- TISSUE SPECIFICITY: SUBSET OF MESODERMAL CELLS.
CC [5]- DISEASE: Defects in TWIST1 are one of the causes of Saethre-
CC Chotzen syndrome (SCS) [MIM:101400]; also known as
CC acrocephalosyndactyly type III (ACS3). SCS is an autosomal
CC dominant defect characterized by minor skull and limb anomalies.
CC [6]- DISEASE: Defects in TWIST1 are the cause of Robinow-Sorauf
CC syndrome (RSS) [MIM:180750]; also known as craniosynostosis-bifid
CC hallux syndrome. RSS is an autosomal dominant defect characterized
CC by minor skull and limb anomalies which is very similar to
CC Saethre-Chotzen syndrome.
CC [7]- DISEASE: Defects in TWIST1 are a cause of Baller-Gerold syndrome
CC (BGS) [MIM:218600]. BGS is a craniosynostosis with radial defects
CC along with other features, including small, round ears with
CC prominent crus helices and cervical anomalies. It seems an
CC heterogeneous form of Saethre-Chotzen syndrome.
CC [8]- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X91662; CAA62850.1; -;
CC EMBL; X99268; CAA67664.1; -;
CC EMBL; U80998; AAC50930.1; -;
CC EMBL; Y10871; CAA71821.1; -;
CC EMBL; AC003986; AAC60381.1; -;
CC EMBL; BC036704; AAH36704.1; -;
CC F01204; G01204;
CC TRANSFAC; T04913; -;
CC Genew; HGNC:12428; TWIST1.
CC MIM; 601622; -;
CC MIM; 101400; -;
CC MIM; 180750; -;
CC MIM; 218600; -;
CC GO; GO:0004857; F:enzyme inhibitor activity; TAS.

DR GO; GO:0005515; P:protein binding activity; TAS.
 DR GO; GO:0007001; P:chromosome organization and biogenesis (seen. . .; TAS.
 DR GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
 DR GO; GO:0000122; P:negative regulation of transcription from P. . .; TAS.
 DR GO; GO:0001501; P:skeletal development; TAS.
 DR InterPro; IPR001092; HLH_basic.
 DR Pfam; PF00010; HLH; 1.
 DR SMART; SM00353; HLH; 1.
 DR PROSITE; PS00038; HLH_1; 1.
 DR PROSITE; PS00888; HLH_2; 1.
 KW Differentiation; Developmental protein; Nuclear protein; DNA-binding;
 KW Transcription regulation; Disease mutation; Polymorphism.
 FT DOMAIN 80 98 GLY-RICH.
 FT DNA BIND 108 120 BASIC DOMAIN.
 FT DOMAIN 121 160 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 FT VARIANT 31 31 S -> G (in dbSNP:1800126).
 FT VARIANT 84 84 G -> S (in dbSNP:2234705).
 FT VARIANT 119 119 Q -> P (in SCS).
 FT VARIANT 131 131 L -> P (in SCS).
 FT VARIANT 135 135 I -> IAALEKII (in SCS).
 FT VARIANT 139 139 P -> PKIITLP (in SCS).
 FT VARIANT 156 156 I -> V (in BGS).
 FT CONFLICT 32 32 G -> A (in REF. 2).
 FT CONFLICT 36 36 G -> A (in REF. 2).
 FT CONFLICT 41 41 S -> T (in REF. 1 AND 4).
 FT CONFLICT 45 45 S -> T (in REF. 1 AND 4).
 FT CONFLICT 56 56 MISSING (in REF. 1 AND 4).
 FT CONFLICT 59 59 G -> A (in REF. 2).
 FT CONFLICT 92 92 G -> GGGG (in REF. 2).
 SQ SEQUENCE 202 AA; 20954 MW; 9394E4351BA1D081 CRC64;
 Query Match 1.9%; Score 7; DB 1; Length 202;
 Best Local Similarity 100.08; Pred. No. 25;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 235 GGGAGGG 241
 Db 84 GGGAGGG 90
 RESULT 22
 TWIS1_MOUSE STANDARD; PRT; 206 AA.
 AC P26687;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Twist related protein 1 (M-twist).
 GN TWIS1 OR TWIST.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91122450; PubMed=1840517;
 RA Wolf C., Thiese C., Stoetzel C., Thiese B., Gerlinger P.,
 RA Perrin-Schmitt F.;
 RT "The M-twist gene of Mus is expressed in subsets of mesodermal cells
 RT and is closely related to the Xenopus X-twi and the Drosophila twist
 RT genes";
 RL Dev. Biol. 143:363-373 (1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N;
 RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Murney D.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
 RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RP FUNCTION.
 RX MEDLINE=98001585; PubMed=9343420;
 RA Hamanori Y., Wu H.Y., Sartorelli V., Kedes L.;
 RT "The basic domain of myogenic basic helix-loop-helix (bHLH) proteins
 RT is the novel target for direct inhibition by another bHLH protein,
 RT Twist";
 RL Mol. Cell. Biol. 17:6563-6573 (1997).
 CC -1- FUNCTION: PROBABLE TRANSCRIPTION FACTOR, WHICH SEEMS TO BE
 CC INVOLVED IN THE NEGATIVE REGULATION OF CELLULAR DETERMINATION AND
 CC IN THE DIFFERENTIATION OF SEVERAL LINEAGES INCLUDING MYOGENESIS,
 CC OSTEOGENESIS, AND NEUROGENESIS. INHIBITS MYOGENESIS BY
 CC SEQUESTERING E PROTEINS, INHIBITING TRANS-ACTIVATION BY MEF2, AND
 CC INHIBITING DNA-BINDING BY MYOD THROUGH PHYSICAL INTERACTION. THIS
 CC INTERACTION PROBABLY INVOLVES THE BASIC DOMAINS OF BOTH PROTEINS.
 CC -1- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
 CC bHLH PROTEIN. HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: SUBSET OF MESODERMAL CELLS.
 CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M63649; AAA0514.1; -;
 CC EMBL; M63650; AAA0515.1; -;
 CC EMBL; BC033434; AAH33434.1; -;
 CC PIR; I53066; I53066.
 CC TRANSPAC; T01635; -;
 CC MGD; MGI:98872; Twist1.
 CC GO; GO:0005634; C:nucleus; NAS.
 CC GO; GO:0003700; F:transcription factor activity; NAS.
 CC GO; GO:0030154; P:cell differentiation; IMP.
 CC GO; GO:0030326; P:limb morphogenesis; IMP.
 CC GO; GO:0045596; P:negative regulation of cell differentiation; IDA.
 CC GO; GO:0045843; P:negative regulation of myogenesis; IDA.
 CC GO; GO:0001679; P:neurulation; IMP.
 CC GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
 CC InterPro; IPR001092; HLH_basic.
 CC Pfam; PF00010; HLH; 1.
 CC SMART; SM00353; HLH; 1.
 CC PROSITE; PS00038; HLH_1; 1.
 CC PROSITE; PS00888; HLH_2; 1.
 KW Differentiation; Developmental protein; Nuclear protein; DNA-binding;
 KW Transcription regulation.
 FT DOMAIN 80 102 GLY-RICH.
 FT DOMAIN 112 124 BASIC DOMAIN.

```

FT DOMAIN      125 164 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT VARIANT     36 36 A -> R (IN CDNA).
FT VARIANT     91 91 G -> P (IN CDNA).
SQ SEQUENCE    206 AA; 21198 MW; 618AD8B9B87C555 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GCGAGGG 241
DB 84 GCGAGGG 90

RESULT 23
GLP1_BRANA STANDARD; PRT; 207 AA.
AC P46271;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Germin-like protein 1 precursor.
GN GER1.
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRAIN=cv. Samourai; Tissue=Seedling;
RX MEDLINE=98009982; PubMed=9349269;
RA Membré N., Berra A., Neutelings G., David A., David H., Staiger D.,
RA Saez Vaquez J., Raynal M., Delaeny M., Bernier P.;
RT "cDNA sequence, genomic organization and differential expression of
RT three Arabidopsis genes for germin/oxalate oxidase-like proteins.";
RL Plant Mol. Biol. 35:459-469(1997).
CC -!- FUNCTION: MAY PLAY A ROLE IN PLANT DEFENSE. HAS PROBABLY NO
CC OXALATE OXIDASE ACTIVITY EVEN IF THE ACTIVE SITE IS CONSERVED.
CC -!- SUBUNIT: OLIGOMER (BELIEVED TO BE A PENTAMER BUT PROBABLY
CC HEXAMER) (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Apoplast (By similarity).
CC -!- SIMILARITY: BELONGS TO THE GERMIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U21743; AA086365.1; -
CC PIR; T07854; T07854.
CC HSSP; P45850; 1P12.
CC InterPro; IPR006045; Cupin.
CC InterPro; IPR007113; Cupin_sup.
CC InterPro; IPR001929; Germin.
CC Pfam; PF00190; Cupin; 1.
CC PRINTS; PR00325; GERMIN.
CC PROSITE; PS00725; GERMIN; 1.
KW Apoplast; Cell wall; Signal; Glycoprotein; Manganese; Metal-binding.
FT SIGNAL      1 17 POTENTIAL.
FT CHAIN       18 207 GERMIN-LIKE PROTEIN 1.
FT METAL       99 99 MANGANESE (BY SIMILARITY).
FT METAL      101 101 MANGANESE (BY SIMILARITY).
FT METAL      106 106 MANGANESE (BY SIMILARITY).
FT METAL      145 145 MANGANESE (BY SIMILARITY).
FT DISULFID    23 38 BY SIMILARITY.
FT CARBOHYD    58 58 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE    207 AA; 21514 MW; B6A29465A73B956E CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GCGAGGG 241
DB 84 GCGAGGG 90

RESULT 24
RR3_CHLVU STANDARD; PRT; 231 AA.
AC P56365;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S3.
GN RPS3.
OS Chlorella vulgaris.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OC Chlorellaceae; Chlorella.
OX NCBI_TaxID=3077;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IAM C-27 / Tamiya;
RX MEDLINE=97303241; PubMed=9159184;
RA Wakasugi T., Nagai T., Kapoor M., Sugita M., Ito M., Ito S.,
RA Tsudzuki J., Nakashima K., Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,
RA Inamura A., Yoshinaga K., Sugita M.;
RT "Complete nucleotide sequence of the chloroplast genome from the
RT green alga Chlorella vulgaris: the existence of genes possibly
RT involved in chloroplast division.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).
CC -!- SUBUNIT: Part of the 30S ribosomal subunit.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- SIMILARITY: BELONGS TO THE S3P FAMILY OF RIBOSOMAL PROTEINS.
CC -!- SIMILARITY: Contains 1 KH type-2 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB001684; BAA58006.1; -
CC PIR; T07358; T07358.
CC HAMAP; MF_01309; -; 1.
CC InterPro; IPR004044; KH TYPE 2.
CC InterPro; IPR001351; Ribosomal_S3.
CC InterPro; IPR005704; S3_bact.
CC Pfam; PF00189; Ribosomal_S3_C; 1.
CC Pfam; PF00417; Ribosomal_S3_N; 1.
CC TIGRPFAMs; TIGR01009; rpsC_bact; 1.
CC PROSITE; PS50823; KH TYPE 2; 1.
CC PROSITE; PS00548; RIBOSOMAL_S3; 1.
KW Ribosomal protein; RNA-binding; rRNA-binding; Chloroplast.
FT DOMAIN     39 123 KH TYPE-2
SQ SEQUENCE    231 AA; 26404 MW; D3EDE90A53551097 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 VQALRA 171
DB 154 VQALRA 160

RESULT 25
CD8A_FELCA STANDARD; PRT; 239 AA.
ID CD8A_FELCA

```

```

Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TVKTLKG 10
DB 195 TVKTLKG 201

RESULT 24
RR3_CHLVU STANDARD; PRT; 231 AA.
AC P56365;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S3.
GN RPS3.
OS Chlorella vulgaris.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OC Chlorellaceae; Chlorella.
OX NCBI_TaxID=3077;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IAM C-27 / Tamiya;
RX MEDLINE=97303241; PubMed=9159184;
RA Wakasugi T., Nagai T., Kapoor M., Sugita M., Ito M., Ito S.,
RA Tsudzuki J., Nakashima K., Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,
RA Inamura A., Yoshinaga K., Sugita M.;
RT "Complete nucleotide sequence of the chloroplast genome from the
RT green alga Chlorella vulgaris: the existence of genes possibly
RT involved in chloroplast division.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).
CC -!- SUBUNIT: Part of the 30S ribosomal subunit.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- SIMILARITY: BELONGS TO THE S3P FAMILY OF RIBOSOMAL PROTEINS.
CC -!- SIMILARITY: Contains 1 KH type-2 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB001684; BAA58006.1; -
CC PIR; T07358; T07358.
CC HAMAP; MF_01309; -; 1.
CC InterPro; IPR004044; KH TYPE 2.
CC InterPro; IPR001351; Ribosomal_S3.
CC InterPro; IPR005704; S3_bact.
CC Pfam; PF00189; Ribosomal_S3_C; 1.
CC Pfam; PF00417; Ribosomal_S3_N; 1.
CC TIGRPFAMs; TIGR01009; rpsC_bact; 1.
CC PROSITE; PS50823; KH TYPE 2; 1.
CC PROSITE; PS00548; RIBOSOMAL_S3; 1.
KW Ribosomal protein; RNA-binding; rRNA-binding; Chloroplast.
FT DOMAIN     39 123 KH TYPE-2
SQ SEQUENCE    231 AA; 26404 MW; D3EDE90A53551097 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 VQALRA 171
DB 154 VQALRA 160

RESULT 25
CD8A_FELCA STANDARD; PRT; 239 AA.
ID CD8A_FELCA

```


P41688;
01-NOV-1995 (Rel. 32, Created)
01-NOV-1995 (Rel. 32, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
T-cell surface glycoprotein CD8 alpha chain precursor.
CD8A.
ON Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Thymus;
RC MEDLINE=94178799; PubMed=8132208;
RA Pecoraro M., Kawaguchi Y., Miyazawa T., Norimine J., Maeda K.,
RA Toyosaki T., Tohya Y., Kai C., Mikami T.;
RT "Isolation, sequence and expression of a cDNA encoding the
RT alpha-chain of the feline CD8";
RL Immunology 81:127-131(1994).
[2]
RN SEQUENCE FROM N.A.
RP TISSUE=Thymus;
RC Miyazawa T.;
RA Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: IDENTIFIES CYTOTOXIC/SUPPRESSOR T-CELLS THAT INTERACT
CC WITH MHC CLASS I BEARING TARGETS. CD8 IS THOUGHT TO PLAY A ROLE IN
CC THE PROCESS OF T-CELL MEDIATED KILLING. CD8 ALPHA CHAINS BINDS TO
CC CLASS I MHC MOLECULES ALPHA-3 DOMAINS.
CC -!- SUBUNIT: IN GENERAL HETERODIMER OF AN ALPHA AND A BETA CHAIN
CC LINKED BY TWO DISULFIDE BONDS. CAN ALSO FORM HOMODIMERS.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D16536; BAA03973.1; -.
CC EMBL; AB000485; BAA19126.1; -.
CC PIR; I46082; I46082.
CC HSP; P01732; 1CD8.
CC GO; GO:0042101; C:T-cell receptor complex; ISS.
CC GO; GO:0015026; F:coreceptor activity; ISS.
CC GO; GO:0042288; F:MHC class I protein binding activity; ISS.
CC GO; GO:0005515; F:protein binding activity; ISS.
CC GO; GO:0006955; P:immune response; ISS.
CC GO; GO:0042110; P:T-cell activation; ISS.
CC GO; GO:0007189; P:transmembrane receptor protein tyrosine kin. . .; ISS.
CC InterPro; IPR007110; IG-like.
CC InterPro; IPR003599; IG.
CC Pfam; PF00047; Ig_1.
CC SMART; SM00409; IG_1.
CC PROSITE; PS0835; IG LIKE; 1.
CC Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW Immune response; Signal.
FT SIGNAL 1 21
FT CHAIN 22 239
FT T-CELL SURFACE GLYCOPROTEIN CD8 ALPHA
FT CHAIN.
FT DOMAIN 22 188
FT TRANSMEM 189 210
FT POTENTIAL.
FT DOMAIN 211 239
FT CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 139
FT IG-LIKE V-TYPE.
FT DISULFID 46 119
FT BY SIMILARITY.
SQ SEQUENCE 239 AA; 26120 MW; F9A171820E8FF1BC CRC64;

Query Match 1.9%; Score 7; DB 1; Length 239;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 APITTSQ 120
Db 153 APITTSQ 159

RESULT 26
HMXU MUSDO STANDARD; PRT; 248 AA.
AC Q254E1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeotic Ultrabithorax protein (Fragment).
GN UBX.
OS Musca domestica (House fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Muscoidea;
OC Muscidae; Musca.
OX NCBI_TaxID=7370;
[1]
RN SEQUENCE FROM N.A.
RP Wilde C.D., Akam M.;
RT "Conserved sequence elements in the 5' region of the Ultrabithorax
RT transcription unit";
RL EMBO J. 6:1393-1401(1987).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC Binds the consensus region 5'-TTAAT[GT][GA]-3' (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE ANTP HOMEBOX FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X05178; CAA28813.1; -.
CC InterPro; IPR001827; Antennapedia.
CC InterPro; IPR001356; Homeobox.
CC PROSITE; PS00032; ANTENNAPEDIA; 1.
CC PROSITE; PS00027; HOMEBOX_1; PARTIAL.
KW Transcription regulation; Activator; Homeobox; DNA-binding;
KW Developmental protein; Nuclear protein.
FT NON TER 1 1
FT DOMAIN 32 38 POLY-ALA.
FT DOMAIN 108 123 POLY-GLY.
FT DOMAIN 142 151 POLY-GLY.
FT DOMAIN 224 227 POLY-ALA.
FT SITE 241 246 ANTP-TYPE HEXAPEPTIDE.
SQ SEQUENCE 248 AA; 23912 MW; AD7CE98CC39CCF2E CRC64;

Query Match 1.9%; Score 7; DB 1; Length 248;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GCGAGGG 241
Db 108 GCGAGGG 114

RESULT 27
GSHI HUMAN STANDARD; PRT; 264 AA.
ID GSHI HUMAN
AC Q9H4S2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein GSH-1.


```

CC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10360;
RN [1]
RX MEDLINE=90269039; PubMed=2161319;
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
RA Horsnell T., Hutchinson C.A. III, Kouzarides T., Martignetti J.A.,
RA "Piedra E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RA "Analysis of the protein-coding content of the sequence of human
RT cytomegalovirus strain AD169."
RL Curr. Top. Microbiol. Immunol. 154:125-169 (1990).
RN [2]
RN SEQUENCE FROM N.A.
RP Wang S.K., Duh C.Y.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUP TOGETHER HSV-6 AND
CC HSV-7 U79 AND HCMV UL112 (P34).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X17403; CAA35315.1; ALT_TERM.
DR EMBL; U57433; AAB53252.1; -.
DR PIR; S09860; S09860.
DR InterPro; IPR004138; U79_P34.
DR Pfam; PF03064; U79_P34; 1.
KW Early protein; Phosphorylation.
FT DOMAIN 148 153 POLY-GLY.
FT DOMAIN 197 200 POLY-SER.
FT DOMAIN 202 220 POLY-GLY.
SQ SEQUENCE 268 AA; 28142 MW; BD5906F4EAB28FC9 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 268;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 GGGAGGG 241
Db 206 GGGAGGG 212
|||||

RESULT 30
CDK4_PIG STANDARD; PRT; 303 AA.
AC P79432;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Cell division protein kinase 4 (EC 2.7.1.-) (Cyclin-dependent kinase
DE 4).
GN CDK4.
OS Sus scrofa (Pig)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RN SEQUENCE FROM N.A.
RP Shibuya H., Renshaw F.G., Bouchard G.F., Nonneman D., Johnson G.S.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
CC -1- ENZYME REGULATION: PHOSPHORYLATION AT T-172 IS NECESSARY FOR
CC ENZYMACTIC ACTIVITY (BY SIMILARITY).
CC -1- SUBUNIT: FORMS A STABLE COMPLEX WITH D-TYPE G1 CYCLINS (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U68478; AAB39636.1; -.
DR EMBL; U68479; AAB39637.1; -.
DR HSSP; Q00534; 1B18.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_SF; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Cell cycle; Cell division; Phosphorylation.
FT DOMAIN 6 295
FT NP_BIND 12 20 ATP (BY SIMILARITY).
FT BINDING 35 35 ATP (BY SIMILARITY).
FT ACT_SITE 140 140 BY SIMILARITY.
FT MOD_RES 172 172 PHOSPHORYLATION (BY CAX) (BY SIMILARITY).
SQ SEQUENCE 303 AA; 33690 MW; D77C5630B25EF20B CRC64;

Query Match 1.9%; Score 7; DB 1; Length 303;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 GGGAGGG 241
Db 42 GGGAGGG 48
|||||

RESULT 31
RALLY_HUMAN STANDARD; PRT; 306 AA.
ID RALLY_HUMAN
AC Q9UKM9; Q14621; Q9BQX6; Q9UJE3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE RNA-binding protein Raly (hnRNP associated with lethal yellow homolog)
DE (Autoantigen p542).
GN RALLY OR P452.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A. (ISOFORM 2).
RP TISSUE=Testis;
RX MEDLINE=99431566; PubMed=10500250;
RA Khrebukova I., Kuklin A., Woychik R.P., Michaud E.J.;
RT "Alternative processing of the human and mouse raly genes.";
RL Blochim. Biophys. Acta 1447:107-112 (1999).
RN [2]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RP Vaughan J.H.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barclay K.F., Bates K.N., Bead L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,

```

RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.B., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosht K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Leharvalaiho M.H., Leveraha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McComachie L.J., McIay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whithead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871 (2001).
RN [4]
RP SEQUENCE OF 85-306 FROM N.A. (ISOFORM 1), AND AUTOIMMUNE DISEASE.
RC TISSUE=Lymphocytes;
RX MEDLINE=98018738; PubMed=9376072;
RA Rhodes G.H., Valbracht J.R., Nguyen M.-D., Vaughan J.H.;
RT "The p542 gene encodes an autoantigen that cross-reacts with EBNA-1 of
RT the Epstein Barr virus and which may be a heterogeneous nuclear
RT ribonucleoprotein.";
RL J. Autoimmun. 10:447-454 (1997).
RN [5]
RP SEQUENCE OF 227-253, AND DETERMINATION OF AUTOANTIGENIC EPITOPE.
RX MEDLINE=95190029; PubMed=7533788;
RA Vaughan J.H., Valbracht J.R., Nguyen M.-D., Handley H.H., Smith R.S.,
RA Patrick K., Rhodes G.H.;
RT "Epstein-Barr virus-induced autoimmune responses. I. Immunoglobulin M
RT autoantibodies to proteins mimicking and not mimicking Epstein-Barr
RT virus nuclear antigen-1.";
RL J. Clin. Invest. 95:1306-1315 (1995).
CC -!- FUNCTION: Probable-RNA binding protein. Could be a heterogeneous
CC nuclear ribonucleoprotein (hnRNP).
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=2;
CC IsoId=Q9UKM9-1; Sequence=Displayed;
CC Name=1;
CC IsoId=Q9UKM9-2; Sequence=VSP_005804;
CC -!- TISSUE SPECIFICITY: Expressed in heart, brain, lung, liver,
CC skeletal muscle, kidney and pancreas. Weakly expressed in
CC placenta.
CC -!- DISEASE: Autoantigen found in infectious mononucleosis caused by
CC Epstein-Barr virus. An epitope recognized by B-cells, which cross-
CC react with the BKRFL protein (EBNA-1 nuclear protein) of Epstein-
CC Barr virus has been identified.
CC -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -!- CAUTION: Ref.4 (CAC29371) sequence differs from that shown due to
CC erroneous gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AF148457; AAF04487.1; -;
CC DR EMBL; L38696; AAC28898.1; -;
CC DR EMBL; AL031668; CAC29371.1; ALT_SEQ.
CC DR EMBL; AL031668; CAB43742.1; -;
CC DR Genew; HGNC:15921; RALY.
CC DR InterPro; IPR000504; RNA_rec_mot.
CC DR Pfam; PF00076; rrm; 1.
CC DR SMART; SM00360; RRM; 1.
CC DR PROSITE; PS50102; RRM; 1.

DR PROSITE; PS00030; RRM_RNP_1; 1.
KW Ribonucleoprotein; RNA-binding; Nuclear protein; Antigen;
KW Alternative splicing; Polymorphism.
FT DOMAIN 21 92 RNA-BINDING (RRM).
FT EPI TOPE (RECOGNIZED BY BKRFL ANTIBODIES).
FT POLY-GLY. 227 253
FT MISSING (in isoform 1). 225 251
FT VARSPLIC 110 125
FT VARIANT 215 215 /FTID=VSP_005804.
FT Q -> R (IN dbSNP:3180568).
FT /FTID=VAR_015223.
FT G -> S (IN dbSNP:2281209).
FT /FTID=VAR_015224.
FT EQ -> DE (IN REF. 2).
FT CONFLICT 214 215
FT CONFLICT 230 230 A -> AS (IN REF. 2).
SQ SEQUENCE 306 AA; 32463 MW; 7F4376D3BD8E4728 CRC64;
Query Match 1.9%; Score 7; DB 1; Length 306;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 235 GGGAGGG 241
Db 227 GGGAGGG 233
|||||
RESULT 32
ID DDL_XYLFA STANDARD; PRT; 320 AA.
AC Q9P79;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE D-alanine-D-alanine ligase (EC 6.3.2.4) (D-alanylalanine synthetase)
DE (D-Ala-D-Ala ligase).
GN DDL OR XF0799.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=2371;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9a5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Coutinho L.L., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Facinani A.P., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Fraga J.S., Ferreira A.J.S., Ferreira V.C.A., Ferto J.A.,
RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hohnsbeil J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.B., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Mend C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A.J., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pequerio J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V. de Rosa A.J.M.,
RA de Rosa V.B. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159 (2000).
CC -!- FUNCTION: Cell wall formation (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + 2 D-alanine = ADP + phosphate + D-

CC alanyl-D-alanine.
CC -1- PATHWAY: D-alanine branch of peptidoglycan biosynthesis; second
CC step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the D-alanine--D-alanine ligase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE003919; AAF83609.1; -
CC PIR; F82763; F82763.
CC HSP; P07862; LIOV.
CC HAMAP; MF_00047; -; 1.
CC InterPro; IPR005905; D_alad_D_al.
CC InterPro; IPR000291; Dala_lig_Van.
CC Pfam; PF01820; Dala_Dala_ligase; 1.
CC TIGRfams; TIGR01205; D_alad_alatIGR; 1.
CC PROSITE; PS00843; DALA_DALA_LIGASE_1; FALSE_NEG.
CC PROSITE; PS00844; DALA_DALA_LIGASE_2; 1.
CC KW Ligase; Cell wall; Peptidoglycan synthesis; Complete proteome.
CC SQ SEQUENCE 320 AA; 34626 MW; DBA8D185318F3FA0 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 320;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 LLSGRNV 145
Db 36 LLSGRNV 42
|||||
|||||

RESULT 33
HXDB MOUSE
ID HXDB MOUSE STANDARD; PRT; 323 AA.
AC P23813;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-D11 (Hox-4.6) (Hox-5.5).
GN HOXD11 OR HOXD-11 OR HOX-4.6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
MEDLINE=91293104; PubMed=1676674;
RA Izpisua-Belmonte J.-C., Falkenstein H., Dolle P., Renucci A.,
RA Duboule D.;
RA "Murine genes related to the Drosophila AbdB homeotic genes are
RT sequentially expressed during development of the posterior part of
RT the body";
RL EMO J. 10:2279-2289(1991).
[2]
SEQUENCE FROM N.A.
STRAIN=129/SV;
RC STRAIN=129/SV;
RX MEDLINE=94074553; PubMed=7902810;
RA Gerard M., Duboule D., Zakany J.;
RA "Structure and activity of regulatory elements involved in the
RT activation of the Hoxd-11 gene during late gastrulation.";
RL EMO J. 12:3539-3550(1993).
[3]
SEQUENCE FROM N.A.
STRAIN=129/SV; TISSUE=Liver;
RC Takada S., Cook M., Kramlauf R., McMahon A.P.;
RL Submitted (MAY-1991) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH

CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING DEVELOPMENT OF THE
CC POSTERIOR PART OF THE BODY.
CC -1- SIMILARITY: BELONGS TO THE ABD-B HOMEBOX FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X60761; CAA43173.1; -
CC EMBL; X60762; CAA43173.1; JOINED.
CC EMBL; X71422; CAA50553.1; -
CC EMBL; X60395; CAA42943.1; ALT_INIT.
CC PIR; S16318; S16318.
CC HSP; P14653; LB72.
CC TRANSFAC; T01760; -
CC GGI; MGI:96203; Hoxd11.
CC GO; GO:0001501; P.skeletal development; IMP.
CC InterPro; IPR001356; Homeobox.
CC Pfam; PF00046; homeobox; 1.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX_1; 1.
CC PROSITE; PS00071; HOMEBOX_2; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT DOMAIN 64 87 POLY-GLY.
FT DOMAIN 93 105 POLY-ALA.
FT DNA_BIND 251 310 HOMEBOX.
SQ SEQUENCE 323 AA; 33545 MW; B003D1347F5FE771 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 323;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
Db 67 GGGAGGG 73
|||||
|||||

RESULT 34
A32C DROME
ID A32C DROME STANDARD; PRT; 329 AA.
AC O46203; O46225; Q8T4D3; Q9TY45; Q9VKL2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Accessory gland protein Acp32CD precursor.
GN ACP32CD OR G4605.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC STRAIN=Canton-S; TISSUE=Male accessory gland;
RX MEDLINE=98135120; PubMed=9474779;
RA Wolfner M.F., Harada H.A., Bertram M.J., Stelick T.J., Kraus K.W.,
RA Kalb J.M., Lung Y.O., Neubaum D.M., Park M., Tram U.K.;
RT "New genes for male accessory gland proteins in Drosophila
RT melanogaster";
RL Insect Biochem. Mol. Biol. 27:825-834(1997).
[2]
SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Ananatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazie R.G., Champ M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Chery J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foslter C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Murny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-F., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195 (2000).
 RN [3]
 RP REVISIONS.
 RC STRAIN=Berkeley;
 RX MEDLINE=22426059; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kamikorn J.S., Milburn G.H., Prochuk S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review."
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley; TISSUE=Testis;
 RX MEDLINE=22426066; PubMed=12537569;
 RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
 RA George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,
 RA Rubin G.M., Celniker S.E.;
 RT "A *Drosophila* full-length cDNA resource."
 RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
 RN [5]
 RP SEQUENCE OF 24-289 FROM N.A.
 RC STRAIN=WS1, WS9, WS12, WS16, WS17, WS19, WS26, WS47, ZIM5, ZIM10,
 RC ZIM22, ZIM24, ZIM30, ZIM32, and ZIM35;
 RX MEDLINE=20556153; PubMed=1102391;
 RA Begun D.J., Whitley P., Todd B.L., Waldrip-Dail H.M., Clark A.G.;
 RT "Molecular population genetics of male accessory gland proteins in
 RT *Drosophila*."
 RL Genetics 156:1879-1888(2000).
 CC -!- FUNCTION: Responsible for physiological and behavioral changes in
 CC mated female flies.
 CC -!- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -!- ALTERNATIVE PRODUCTS:

CC Event-Alternative Initiation;
 CC Comment=2 isoforms, Long (shown here) and Short, are produced by
 CC alternative initiation;
 CC -!- TISSUE SPECIFICITY: Seminal fluid.
 CC -!- CAUTION: Ref.3 and Ref.4 sequences differ from that shown due to
 CC frameshifts. These may be a natural frameshifts and this protein
 CC could be a pseudogene.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U85764; AAB96388.1; -
 CC EMBL; U85764; AAB96389.1; -
 CC EMBL; U90948; AAB96395.1; -
 CC EMBL; AE003630; AAF53055.2; ALT_FRAME.
 CC EMBL; AY089245; AAL89983.1; ALT_FRAME.
 CC EMBL; AY010544; AAG32576.1; -
 CC EMBL; AY010545; AAG32577.2; -
 CC EMBL; AY010546; AAG32578.1; -
 CC EMBL; AY010547; AAG32579.1; -
 CC EMBL; AY010548; AAG32580.1; -
 CC EMBL; AY010549; AAG32581.2; -
 CC EMBL; AY010550; AAG32582.1; -
 CC EMBL; AY010551; AAG32583.1; -
 CC EMBL; AY010552; AAG32584.1; -
 CC EMBL; AY010553; AAG32585.1; -
 CC EMBL; AY010554; AAG32586.1; -
 CC EMBL; AY010555; AAG32587.1; -
 CC EMBL; AY010556; AAG32588.1; -
 CC EMBL; AY010557; AAG32589.1; -
 CC EMBL; AY010558; AAG32590.1; -
 CC EMBL; AY010559; AAG32591.1; -
 CC FlyBase; FBgn0023415; ACP32CD.
 CC GO; GO:0005576; C:extracellular; IE.
 CC GO; GO:0005180; F:peptide hormone; NAS.
 CC GO; GO:0045434; P:negative regulation of female receptivity, . . .; NAS.
 KW Alternative Initiation; Signal; Behavior; Polymorphism.
 FT SIGNAL 1 21
 FT CHAIN 22 329
 FT ACCESSORY GLAND PROTEIN ACP32CD, ISOFORM
 FT LONG.
 FT ACCESSORY GLAND PROTEIN ACP32CD, ISOFORM
 FT SHORT.
 FT FOR ISOFORM SHORT.
 FT INIT MET 89 89
 FT DOMAIN 100 182
 FT DOMAIN 200 326
 FT VARIANT 45 45
 FT D -> E (IN STRAINS WS1, WS9, WS16, ZIM5,
 FT ZIM10, ZIM24, ZIM30, ZIM32 AND
 FT ZIM35).
 FT A -> D (IN STRAINS 7, WS1, WS12, WS16,
 FT WS17, ZIM10, ZIM22, ZIM32 AND
 FT ZIM35).
 FT A -> T (IN STRAINS 7, WS1, WS9, WS16,
 FT WS17 AND ZIM10).
 FT L -> P (IN STRAINS WS9 AND WS19).
 FT TRM -> PRT (IN STRAINS WS9, ZIM30 AND
 FT ZIM22).
 FT M -> T (IN STRAINS WS12, WS17 AND
 FT WS26).
 FT SEQUENCE 329 AA; 35557 MW; 5655CCA73759CFA CRC64;
 SQ
 Query Match 1.9%; Score 7; DB 1; Length 329;
 Best Local Similarity 100.0%; Pred.No. 42;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 235 GGGAGGG 241
 DB 110 GGGAGGG 116

```
RESULT 35
TAL_MOUSE
ID TAL_MOUSE STANDARD; PRT; 329 AA.
AC P22091;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE T-cell acute lymphocytic leukemia-1 protein homolog (TAL-1 protein)
DE (Stem cell protein).
DE TAL1 OR TAL-1 OR SCL.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Bone marrow macrophage;
RA MEDLINE=91126096; PubMed=1704135;
RA Begley C.G., Visvader J., Green A.R., Aplan P.D., Metcalf D.,
RA Kirsch I.R., Gough N.M.;
RA "Molecular cloning and chromosomal localization of the murine homolog
RT of the human helix-loop-helix gene SCL.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:869-873 (1991).
CC -1- FUNCTION: IMPLICATED IN THE GENESIS OF HEMOPOIETIC MALIGNANCIES.
CC IT MAY PLAY AN IMPORTANT ROLE IN HEMOPOIETIC DIFFERENTIATION.
CC SERVES AS A POSITIVE REGULATOR OF ERYTHROID DIFFERENTIATION.
CC -1- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION.
CC BHLH PROTEIN. FORMS HETERODIMERS WITH E12 AND E47. BINDS TO THE
CC LIM DOMAIN CONTAINING PROTEIN RHOMBOTIN-2 AND TO DRG1.
CC -1- TISSUE SPECIFICITY: ERYTHROID AND MYELOID CELLS.
CC -1- DOMAIN: THE HELIX-LOOP-HELIX DOMAIN IS NECESSARY AND SUFFICIENT
CC FOR THE INTERACTION WITH DRG1.
CC -1- PTM: PHOSPHORYLATED ON SERINE RESIDUES (BY SIMILARITY).
CC -1- DISEASE: INVOLVED IN CHROMOSOMAL TRANSLOCATION IN LEUKEMIC
CC STEM-CELLS.
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M59764; AAA40097.1; -.
DR EMBL; U01530; AAA86937.1; -.
DR PIR; A37864; A37864.
DR TRNSPAC; T01799; -.
DR MGD; MGI:98480; Tal1.
DR GO; GO:0030218; P:erythrocyte differentiation; IDA.
DR GO; GO:0030097; P:hemopoiesis; IDA.
DR InterPro; IPR001092; HLH_bas.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00038; HLH_1; 1.
DR PROSITE; PS50888; HLH_2; 1.
DR DNA-binding; Transcription regulation; Phosphorylation.
KW Proto-oncogene; Chromosomal translocation.
FT DNA BIND 187 199
FT DOMAIN 200 240
FT DOMAIN 263 272
FT POLY-GLY.
SQ SEQUENCE 329 AA; 34278 MW; 189480B6993CB371 CRC64;
Query Match 1.9%; Score 7; DB 1; Length 329;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 235 GGGAGGG 241
|||
DB 266 GGGAGGG 272
```

```
RESULT 36
SIX3_HUMAN
ID SIX3_HUMAN STANDARD; PRT; 332 AA.
AC G95343;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Homeobox protein SIX3 (Sine oculis homeobox homolog 3).
GN SIX3.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99107815; PubMed=9889003;
RA Granadino B., Gallardo M.E., Lopez-Rios J., Sanz R., Ramos C.,
RA Ayuso C., Bovolenta P., Rodriguez de Cordoba S.;
RA "Genomic cloning, structure, expression pattern, and chromosomal
RT location of the human SIX3 gene.";
RL Genomics 55:100-105 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99377859; PubMed=10454822;
RA Leppert G.S., Yang J.-M., Sundin O.H.;
RA "Sequence and location of SIX3, a homeobox gene expressed in the human
RT eye.";
RL Ophthalmic Genet. 20:1-15 (1999).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Retina;
RA Clark B.J., Hanson I.M., Brown A.G., Ferrier R.K., Prosser J.,
RA van Heyningen V.;
RA "SIX3, a member of the Sine oculis/Six family of transcription
RT factors, is expressed in the developing and adult human eye.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP VARIANTS HPE2 VAL-226; ALA-250 AND PRO-257.
RX MEDLINE=99295940; PubMed=10369266;
RA Wallis D.E., Roessler E., Hehr U., Nanni L., Wiltshire T.,
RA Richter-Costa A., Gillespie-Kaesbach G., Zackai E.H., Rommens J.,
RA Muenke M.;
RA "Mutations in the homeodomain of the human SIX3 gene cause
RT holoprosencephaly";
RL Nat. Genet. 22:196-198 (1999).
CC -1- FUNCTION: MAY BE INVOLVED IN VISUAL SYSTEM DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DISEASE: Defects in SIX3 are the cause of holoprosencephaly type 2
CC (HPE2) [MIM:157170]. HPE2 is a common, severe malformation of the
CC brain that involves separation of the central nervous system into
CC left and right halves.
CC -1- SIMILARITY: BELONGS TO THE SIX3/SINE OCULIS HOMEBOX FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF092047; AAD11939.1; -.
DR EMBL; AF049339; AAD15753.1; -.
DR EMBL; AF083891; AAD51091.1; -.
DR EMBL; AJ012611; CAB42539.1; -.
DR HSSP; P41778; 1DU6.
DR TRNSPAC; T03282; -.
DR Genew; HGNC:10889; SIX3.
DR MIN; 603714; -.
DR MIN; 157170; -.
DR GO; GO:0007420; P:brain development; TAS.
```

```

DR GO: 0007601; P: vision; TAS.
DR InterPro: IPR001356; Homeobox.
DR InterPro: IPR000047; HTH_lambrepreser.
DR InterPro: IPR007105; SIX.
DR InterPro: IPR007106; SIX_SINE_homeo.
DR Pfam: PF000046; homeobox; 1.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX 1; FALSE_NEG.
DR PROSITE; PS00071; HOMEBOX 2; 1.
DR Developmental protein; Homeobox; DNA-binding; Nuclear protein;
KW Disease mutation; Holoprosencephaly.
FT DOMAIN 33 69 POLY-GLY.
FT DOMAIN 206 265 HOMEBOX.
FT DOMAIN 263 266 POLY-ALA.
FT VARIANT 226 226 L -> V (in HPE2).
FT VARIANT 250 250 /FTId=VAR_003771.
FT VARIANT 257 257 V -> A (in HPE2).
FT VARIANT 257 257 /FTId=VAR_003772.
FT VARIANT 257 257 R -> P (in HPE2).
FT VARIANT 257 257 /FTId=VAR_003773.
SQ SEQUENCE 332 AA; 35486 MW; 21EA07F6A2DD978F CRC64;

Query Match 1.9%; Score 7; DB 1; Length 332;
Best Local Similarity 100.0%; Pred.No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
DB 41 GGGAGGG 47

RESULT 37
SIX3 MOUSE STANDARD; PRT; 333 AA.
AC Q62233; P70176; P70177;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein SIX3 (Sine oculis homeobox homolog 3).
GN SIX3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Embryonic brain;
RX MEDLINE=96125147; PubMed=8575305;
RA Oliver G., Mailhos A., Wehr R., Copeland N.G., Jenkins N.A.,
RA Gruss P.;
RT "Six3, a murine homologue of the sine oculis gene, demarcates the
RT most anterior border of the developing neural plate and is expressed
RT during eye development.";
RL Development 121:4045-4055 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=96409319; PubMed=8814301;
RA Kawakami K., Ohno H., Takizawa T., Saito T.;
RT "Identification and expression of Six family genes in mouse retina.";
RL FEBS Lett. 393:259-263 (1996).
CC -1- FUNCTION: May be involved in visual system development.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Six3B;
CC IsoId=Q62233-1; Sequence=Displayed;
CC Name=Six3A;
CC IsoId=Q62233-2; Sequence=VSP_002291, VSP_002292;
CC -1- DEVELOPMENTAL STAGE: FIRST EXPRESSED AT E6.5 OF EMBRYO DEVELOPMENT
CC AROUND THE ANTERIOR BORDER. AT E8.5, EXPRESSION IS FOUND OVER THE

```

```

CC ANTERIOR NEURAL PLATE. AT E9.5, IN THE DIENTEPHALIC PART OF THE
CC VENTRAL FOREBRAIN, OPTIC VESICLES, OLFACTORY PLACODES AND RATHKE'S
CC POUCH. IN LATER STAGES, PRESENT IN HYPOTHALAMUS, EYES AND
CC PITUITARY.
CC -1- SIMILARITY: BELONGS TO THE SIX/SINE OCULIS HOMEBOX FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X50871; CAAG2379.1; ALT_INIT.
CC EMBL; D83144; BAA11822.1; -.
CC EMBL; D83145; BAA11823.1; -.
CC HSSP; P41778; 1DU6.
CC TRANSFAC; T03263; -.
CC TRANSFAC; T03270; -.
CC MGD; MGI:102764; Six3.
CC GO: GO:0005634; C:nucleus; IDA.
CC GO: GO:0005515; P:protein binding activity; IPI.
CC GO: GO:0014564; F:transcriptional repressor activity; IDA.
CC GO: GO:0001654; P:eye morphogenesis; IDA.
CC GO: GO:0016481; P:negative regulation of transcription; IDA.
CC InterPro: IPR001356; Homeobox.
CC InterPro: IPR007105; SIX.
CC InterPro: IPR007106; SIX_SINE_homeo.
CC Pfam; PF00046; homeobox; 1.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX 1; FALSE_NEG.
CC PROSITE; PS00071; HOMEBOX 2; 1.
KW Developmental protein; Homeobox; DNA-binding; Nuclear protein;
KW Alternative splicing.
FT DOMAIN 33 70 GLY-RICH.
FT DOMAIN 207 266 HOMEBOX.
FT DOMAIN 264 267 POLY-ALA.
FT VARSPLIC 271 286 LQQAIGPGSMRSLAE -> SVAGTAARPPQA (in
FT isoform SIX3A).
FT /FTId=VSP_002291.
FT Missing (in isoform SIX3A).
FT /FTId=VSP_002292.
FT G -> GG (IN REF. 1).
FT VA -> WP (IN REF. 1).
FT PSMRSLAEPGCPHGSAPSPSTAASPTTSVSLTERADTG
FT TSILSVTSSDSDCV -> ERDALPGARLPHARLSRVTH
FT GGQPDHQCVPDQARGHRHFDPLGNLQRLGM (IN REF.
FT 1).
SQ SEQUENCE 333 AA; 35592 MW; 1AD7D3C4388043B9 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 333;
Best Local Similarity 100.0%; Pred.No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
DB 42 GGGAGGG 48

RESULT 38
SYEA_CLOAB STANDARD; PRT; 339 AA.
AC Q97GK9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phenylalanine-tRNA synthetase alpha chain (EC 6.1.1.20)
DE (Phenylalanine--tRNA ligase alpha chain) (PhRS).
GN PHRS OR CAC2357.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

```


Clostridium.
NCBI_TaxID=1488;
[1]
SEQUENCE FROM N.A.
STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breston G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838 (2001).
CC -1- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) = AMP +
CC diphosphate + L-phenylalanyl-tRNA(Phe).
CC -1- COFACTOR: Binds 2 magnesium ions per tetramer (By similarity).
CC -1- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC Phe-tRNA synthetase alpha chain subfamily 1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AE007736; AAK80313.1; -
DR PIR; F97190; F97190.
DR HAMAP; MF_00281; -; 1.
DR InterPro; IPR004188; Phe_tRNA_synt_N.
DR InterPro; IPR004529; PheS.
DR InterPro; IPR002319; tRNA-synt 2d.
DR InterPro; IPR006195; tRNA_ligase_II.
DR Pfam; PF02912; Phe_tRNA_synt_N; 1.
DR Pfam; PF01409; tRNA-synt_2d; 1.
DR TIGRFAMs; TIGR00468; pheS; 1.
DR PROSITE; PS50862; AA tRNA_LIGASE_II; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Metal-binding; Magnesium; Complete proteome.
FT METAL 254 254 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 339 AA; 38508 MW; 229E76449BE10918 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 339;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 KKNIEEI 33
Db 22 KKNIEEI 28
|||||

RESULT 39
ROAL SCHAM
ID ROAL SCHAM STANDARD; PRT; 342 AA.
AC P21522;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Heterogeneous nuclear ribonucleoprotein A1, A2/B1 homolog.
GN HNRNP.
OS Schistocerca americana (American grasshopper).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Psterygota;
OC Neoptera; Orthopteroidea; Orthoptera; Gaelefera; Acridomorpha;
OC Acridoidea; Acrididae; Cyrtacanthacridinae; Schistocerca.
OX NCBI_TaxID=7009;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91195067; PubMed=1901646;
RA Ball E.E., Rehm E.J., Goodman C.S.;

"Cloning of a grasshopper cDNA coding for a protein homologous to the
RT A1, A2/B1 proteins of mammalian hnRNP.";
RL Nucleic Acids Res. 19:397-397(1991).
CC -1- FUNCTION: THIS PROTEIN IS A COMPONENT OF RIBONUCLEOSOMES.
CC -1- DEVELOPMENTAL STAGE: ABUNDANT IN EMBRYOS YOUNGER THAN 50% OF
CC EMBRYONIC DEVELOPMENT.
CC -1- SIMILARITY: BELONGS TO THE A/B GROUP OF HNRNP, WHICH ARE BASIC AND
CC GLY-RICH PROTEINS.
CC -1- SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X54670; CAA38481.1; -
DR PIR; S14432; S14432.
DR HGSP; P09651; IUP1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS0102; RRM; 2.
DR PROSITE; PS00030; RRM_RNP_1; 2.
KW Nuclear protein; RNA-binding; Ribonucleoprotein; Repeat.
FT DOMAIN 17 93 RNA-BINDING (RRM) 1.
FT DOMAIN 108 185 RNA-BINDING (RRM) 2.
FT DOMAIN 187 339 GLY-RICH.
SQ SEQUENCE 342 AA; 35404 MW; F8426A21C8125DD8 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 342;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
Db 202 GGGAGGG 208
|||||

RESULT 40
BET3 MESAU
ID BET3 MESAU STANDARD; PRT; 367 AA.
AC O09029; 2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE BET3 protein.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96140430; PubMed=9552091;
RA Peyton M., Stellrecht C.M.M., Naya F.J., Huang H.-P., Samora P.J.,
RA Tsai M.-J.;
RT "BET3, a novel helix-loop-helix protein, can act as a negative
RT regulator of BET2 and MyoD-responsive genes.";
RL Mol. Cell. Biol. 16:626-633(1996).
CC -1- FUNCTION: INHIBITS DNA BINDING OF TCF3 (B47) HOMODIMERS AND TCF3
CC (B47) / NEUROD1 HETERODIMERS AND ACTS AS A STRONG REPRESSOR OF
CC NEUROD1 AND MYOD-RESPONSIVE GENES, PROBABLY BY HETERODIMERIZATION
CC WITH CLASS A BASIC HELIX-LOOP-HELIX FACTORS. DESPITE THE PRESENCE
CC OF AN INTACT BASIC DOMAIN, DOES NOT BIND TO DNA.
CC -1- SUBUNIT: HETERODIMER WITH OTHER BHLH PROTEINS, LIKE TCF3 (B47).
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- TISSUE SPECIFICITY: KIDNEY, LUNG, BRAIN AND PANCREAS (INSULINOMA).
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS. "ATONAL" SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; S80870; AAB50691.1; -;
 DR TRANSFAC; T01674; -;
 DR InterPro; IPR001092; HLH_basic.
 DR Pfam; PF00010; HLH; 1.
 DR SMART; SM00353; HLH; 1.
 DR PROSITE; PS00038; HLH; 1; 1.
 DR PROSITE; PS00888; HLH; 2; 1.
 KW Nuclear protein; Transcription regulation; Repressor.
 FT DOMAIN 11 14 POLY-ALA.
 FT DOMAIN 58 62 POLY-SER.
 FT DOMAIN 83 99 POLY-GLY.
 FT DOMAIN 174 179 POLY-GLY.
 FT DOMAIN 204 217 POLY-GLY.
 FT DNA BIND 229 240 BASIC DOMAIN.
 FT DOMAIN 241 282 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 FT DOMAIN 311 319 POLY-ALA.
 SQ SEQUENCE 367 AA; 35905 MW; 6CAB9AFF96B585F77 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 367;

Best Local Similarity 100.0%; Pred. No. 47;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241

DB 85 GGGAGGG 91

RESULT 41

NLPD_SALTI
 ID NLPD_SALTI STANDARD; PRT; 373 AA.
 AC Q56131;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Lipoprotein nlpd precursor.
 GN NLPD OR STV3050 OR T2826.
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagels K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrall B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ty2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;
 RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
 RT and CT18.";
 RL J. Bacteriol. 185:2330-2337(2003).
 RN [3]

RP SEQUENCE OF 255-373 FROM N.A.
 RC STRAIN=Ty21a;
 RX MEDLINE=95220644; PubMed=7705608;
 RA Robbe-Saule V., Coyneault C., Norel F.;
 RT "The live oral typhoid vaccine Ty21a is a rps mutant and is
 RT susceptible to various environmental stresses.";
 RL FEMS Microbiol. Lett. 126:171-176(1995).
 RN [4]
 RP SEQUENCE OF 269-373 FROM N.A.
 RC STRAIN=Ty2 / ATCC 700931;
 RX MEDLINE=99118315; PubMed=9919662;
 RA Robbe-Saule V., Norel F.;
 RT "The rps mutant allele of Salmonella typhi Ty2 is identical to that
 RT of the live typhoid vaccine Ty21a.";
 RL FEMS Microbiol. Lett. 170:141-143(1999).
 CC -|- FUNCTION: MAY BE INVOLVED IN STATIONARY-PHASE SURVIVAL.
 CC -|- SUBCELLULAR LOCATION: Attached to the inner membrane by a lipid
 CC anchor (Potential).
 CC -|- SIMILARITY: BELONGS TO THE E. COLI NLPD / HAEMOPHILUS LPPB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; AL627276; CAD06031.1; -;
 DR EMBL; AF016843; AA070383.1; -;
 DR EMBL; X81641; CAA57297.1; -;
 DR EMBL; Y17610; CAA76806.1; -;
 DR MEROPS; M37.UFW; -;
 DR InterPro; IPR002482; LysM.
 DR InterPro; IPR002886; Peptidase_M37.
 DR Pfam; PF01476; LysM; 1.
 DR Pfam; PF01551; Peptidase_M37; 1.
 DR SMART; SM00257; LysM; 1.
 DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
 KW Inner membrane; Lipoprotein; Repeat; Signal; Complete proteome.
 FT SIGNAL 1 25 BY SIMILARITY.
 FT CHAIN 26 373 LIPOPROTEIN NLPD.
 FT LIPID 26 26 N-ACYL DIGLYCERIDE (BY SIMILARITY).
 FT DOMAIN 199 246 4 X 7 AA APPROXIMATE REPEATS.
 FT REPEAT 199 205 1.
 FT REPEAT 221 227 2.
 FT REPEAT 233 239 3.
 FT REPEAT 240 246 4.
 SQ SEQUENCE 373 AA; 39183 MW; DD075BB440098FF7 CRC64;
 Query Match 1.9%; Score 7; DB 1; Length 373;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 81 GSTGTSS 87
 DB 344 GSTGTSS 350
 RESULT 42
 ID GLK1 TRIVA STANDARD; PRT; 375 AA.
 AC Q9GTW9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glucokinase 1 (EC 2.7.1.2) (Glucose kinase 1) (Hexokinase 1).
 GN GK1.
 OS Trichomonas vaginalis.
 OC Eukaryota; Parabasalides; Trichomonadida; Trichomonadidae;
 OC Trichomonadinae; Trichomonas.
 OX NCBI_TaxID=5722;
 RN [1]

```

RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.
RC STRAIN=ATCC 30001 / NIH-C1;
RX MEDLINE=21147937; PubMed=11250082;
RA Wu G., Henze K., Mueller M.;
RT "Evolutionary relationships of the glucokinase from the
RL amitochondriate protist, Trichomonas vaginalis.";
CC Gene 264:265-271(2001).
CC -|- CATALYTIC ACTIVITY: ATP + D-glucose = ADP + D-glucose 6-phosphate.
CC -|- SUBUNIT: Monomer.
CC -|- PTM: The N-terminus is blocked.
CC -|- SIMILARITY: BELONGS TO THE BACTERIAL GLUCOKINASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF248652; AAG17616.1; -.
CC DR Transferase; Kinase; Glycolysis; ATP-binding.
CC KW NP_BIND 25 30 ATP (POTENTIAL)
CC FT SEQUENCE 375 AA; 41585 MW; 229F76B24637BE21 CRC64;
CC SQ
CC
CC Query Match 1.9%; Score 7; DB 1; Length 375;
CC Best Local Similarity 100.0%; Pred.No. 48;
CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC Qy 205 TDRAPTQ 211
CC Db 155 TDRAPTQ 161
CC
CC RESULT 43
CC DNAAJ LISIN
CC ID DNAAJ LISIN STANDARD; PRT; 376 AA.
CC AC Q92BN9;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DE 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Chaperone protein dnaJ.
CC GN DNAAJ OR L1N1509.
CC OS Listeria innocua.
CC OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
CC OX NCBI_TaxID=1642;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=CLIP 11262 / Serovar 6a;
CC RX MEDLINE=21537279; PubMed=11679669;
CC RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
CC RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
CC RA Charbit A., Chetouani F., Couve E., de Daruvar A., Deboux P.,
CC RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
CC RA Entian K.-D., Fshih H., Garcia-del Portillo F., Garrido P.,
CC RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
CC RA Jones L.-M., Kaerst U., Krest J., Kuhn M., Kunet F., Kurapkat G.,
CC RA Madueno E., Maitournan A., Mata Vicente J., Ng E., Nedjari H.,
CC RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
CC RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
CC RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
CC RT "Comparative genomics of Listeria species.";
CC RL Science 294:849-852(2001).
CC
CC -|- FUNCTION: ACTS AS A CO-CHAPERONE. STIMULATES, JOINTLY WITH GRPE,
CC THE ATPASE ACTIVITY OF DNAK (BY SIMILARITY).
CC -|- CAPCITOR: BINDS TWO ZINC IONS PER MONOMER (BY SIMILARITY).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- SIMILARITY: BELONGS TO THE DNAAJ FAMILY.
CC -|- SIMILARITY: Contains 1 J domain.
CC -|- SIMILARITY: Contains 1 CR domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL596168; CAC96740.1; -.
CC DR PIR; AD1621; AD1621.
CC DR L1stiliat; L1N01509; -.
CC DR InterPro; IPR002939; DnaJ_C.
CC DR InterPro; IPR001305; DnaJ_CXXCXGKG.
CC DR InterPro; IPR001623; DnaJ_N.
CC DR Pfam; PF00226; DnaJ; 1.
CC DR Pfam; PF01556; DnaJ_C; 1.
CC DR Pfam; PF00684; DnaJ_CXXCXGKG; 1.
CC DR PRINTS; PR00625; DNAAJPROTEIN.
CC DR SMART; SM00271; DnaJ; 1.
CC DR PROSITE; PS00636; DnaJ_1; 1.
CC DR PROSITE; PS00076; DnaJ_2; 1.
CC DR PROSITE; PS00637; DnaJ_CXXCXGKG; 1.
CC KW Chaperone; DNA replication; Heat shock; Repeat; Zinc; Metal-binding;
CC FT DOMAIN 5 69 J-DOMAIN.
CC FT REPEAT 76 117 GLY-RICH.
CC FT REPEAT 146 153 CXXCXGKG MOTIF.
CC FT REPEAT 163 170 CXXCXGKG MOTIF.
CC FT REPEAT 189 196 CXXCXGKG MOTIF.
CC FT REPEAT 203 210 CXXCXGKG MOTIF.
CC FT METAL 146 146 ZINC 1 (BY SIMILARITY).
CC FT METAL 149 149 ZINC 1 (BY SIMILARITY).
CC FT METAL 163 163 ZINC 2 (BY SIMILARITY).
CC FT METAL 166 166 ZINC 2 (BY SIMILARITY).
CC FT METAL 189 189 ZINC 2 (BY SIMILARITY).
CC FT METAL 192 192 ZINC 2 (BY SIMILARITY).
CC FT METAL 203 203 ZINC 1 (BY SIMILARITY).
CC FT METAL 206 206 ZINC 1 (BY SIMILARITY).
CC SQ SEQUENCE 376 AA; 41020 MW; CA8631D6FA120C24 CRC64;
CC
CC Query Match 1.9%; Score 7; DB 1; Length 376;
CC Best Local Similarity 100.0%; Pred.No. 48;
CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC Qy 235 GCGAGGG 241
CC Db 78 GCGAGGG 84
CC
CC RESULT 44
CC DNAAJ LISMO
CC ID DNAAJ LISMO STANDARD; PRT; 377 AA.
CC AC Q955A3;
CC DT 16-OCT-2001 (Rel. 40, Created)
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Chaperone protein dnaJ.
CC GN DNAAJ OR LM01472.
CC OS Listeria monocytogenes
CC OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
CC OX NCBI_TaxID=1639;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=10403S;
CC RX MEDLINE=20163771; PubMed=10701836;
CC RA Hanawa T., Kai M., Kamiya S., Yamamoto T.;
CC RT "Cloning, sequencing, and transcriptional analysis of the dnaK heat
CC shock operon of Listeria monocytogenes.";
CC RL Cell Stress Chaperones 5:21-29(2000).
CC RN [2]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=EGD-e / Serovar 1/2a;
CC RX MEDLINE=21537279; PubMed=11679669;
CC RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,

```

RA Baquero F., Berche P., Bloecher H., Brandt P., Chakraborty T.,
RA Charbit A., Chatouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Duesurget O.,
RA Entian K.-D., Faihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Gobel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krefit J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Matounam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordliek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RT "Comparative genomics of *Listeria species*.";
RL Science 294:849-852(2001).
CC -1- FUNCTION: ACTS AS A CO-CHAPERONE. STIMULATES, JOINTLY WITH GRPE,
CC THE ATPASE ACTIVITY OF DNAK (BY SIMILARITY).
CC -1- COFACTOR: BINDS TWO ZINC IONS PER MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE DNAJ FAMILY.
CC -1- SIMILARITY: Contains 1 CR domain.
CC -1- SIMILARITY: Contains 1 J domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB023064; BAA82790.1; -;
DR EMBL; AL591979; CAC99550.1; -;
DR PIR; AH1258; AH1258.
DR PIR; T43739; T43739.
DR HSP; P08622; LBQZ.
DR ListIlist; LMO01472; -;
DR InterPro; IPR02939; DnaJ_C.
DR InterPro; IPR001305; DnaJ_CXXCXXG.
DR InterPro; IPR001623; DnaJ_N.
DR InterPro; IPR003095; Hsp_DnaJ.
DR Pfam; PF00226; DnaJ; 1.
DR Pfam; PF01556; DnaJ_C; 1.
DR Pfam; PF00684; DnaJ_CXXCXXG; 1.
DR PRINTS; PR00625; DNAJPROTEIN.
DR SMART; SM00271; DnaJ; 1.
DR PROSITE; PS0076; DnaJ_1; 1.
DR PROSITE; PS0076; DnaJ_2; 1.
DR PROSITE; PS00637; DnaJ_CXXCXXG; 1.
KW Chaperone; DNA replication; Heat shock; Repeat; Zinc; Metal-binding;
KW Complete proteome.
FT DOMAIN 5 69 J-DOMAIN.
FT DOMAIN 76 117 GLY-RICH.
FT REPEAT 147 154 CXXCXXG MOTIF.
FT REPEAT 164 171 CXXCXXG MOTIF.
FT REPEAT 190 197 CXXCXXG MOTIF.
FT REPEAT 204 211 CXXCXXG MOTIF.
FT METAL 147 147 ZINC 1 (BY SIMILARITY).
FT METAL 150 150 ZINC 1 (BY SIMILARITY).
FT METAL 164 164 ZINC 2 (BY SIMILARITY).
FT METAL 167 167 ZINC 2 (BY SIMILARITY).
FT METAL 190 190 ZINC 2 (BY SIMILARITY).
FT METAL 193 193 ZINC 2 (BY SIMILARITY).
FT METAL 204 204 ZINC 1 (BY SIMILARITY).
FT METAL 207 207 ZINC 1 (BY SIMILARITY).
SQ SEQUENCE 377 AA; 41077 MW; 70A8C5836239310A CRC64;
Query Match 1.9%; Score 7; DB 1; Length 377;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGG 241
Db 79 GGGAGG 85

RESULT 45
NLPD SALDU
ID NLPD SALDU STANDARD; PRT; 377 AA.
AC P39760;
DT 01-FEB-1995 (Rel. 31, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Lipoprotein nlpd precursor.
GN NLPD.
OS Salmonella dublin.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=98360;
RN [1]
RP SEQUENCE FROM N.A., AND INDUCTION.
RC STRAIN-Lane;
RX MEDLINE=99138748; PubMed=9973354;
RA Paesold G., Krause M.;
RT "Analysis of rpoS mRNA in Salmonella dublin: identification of
RT multiple transcripts with growth-phase dependent variation in
RT transcript stability.";
RL J. Bacteriol. 181:1264-1268(1999).
RN [2]
RP SEQUENCE OF 284-377 FROM N.A.
RC STRAIN-Lane;
RA Krause M.W., El-Gedaily A.;
RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAY BE INVOLVED IN STATIONARY-PHASE SURVIVAL.
CC -1- SUBCELLULAR LOCATION: Attached to the inner membrane by a lipid
CC anchor (Potential).
CC -1- SIMILARITY: BELONGS TO THE E. COLI NLPD / HAEMOPHILUS LPPB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ006131; CAA06881.1; -;
DR EMBL; X82129; CAA57639.1; -;
DR MEROPS; M37.UFW; -;
DR InterPro; IPR002482; LysM.
DR InterPro; IPR002886; Peptidase_M37.
DR Pfam; PF01476; LysM; 1.
DR Pfam; PF01551; Peptidase_M37; 1.
DR SMART; SM00257; LysM; 1.
DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
KW Inner membrane; Lipoprotein; Repeat; Signal.
FT SIGNAL 1 25
FT CHAIN 26 377 LIPOPROTEIN NLPD.
FT LIPID 26 26 N-ACYL DIGLYCERIDE (BY SIMILARITY).
FT DOMAIN 203 250 4 X 7 AA APPROXIMATE REPEATS.
FT REPEAT 203 209 1.
FT REPEAT 225 231 2.
FT REPEAT 237 243 3.
FT REPEAT 244 250 4.
SQ SEQUENCE 377 AA; 39671 MW; DC7F59B943F86512 CRC64;
Query Match 1.9%; Score 7; DB 1; Length 377;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 GSTGTSS 87
Db 348 GSTGTSS 354

RESULT 46
NLPD SALTY
ID NLPD SALTY STANDARD; PRT; 377 AA.
AC P40827;

DT 01-FEB-1995 (Rel. 31, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Lipoprotein nlpD precursor.
GN NLPD OR STM2925
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RL Nature 413:852-856(2001).
RN [2]
RP SEQUENCE OF 280-377 FROM N.A.
RC STRAIN=ATCC 14028S;
RX MEDLINE=94368857; PubMed=8086465;
RA Prince R.W., Fang F.C., Libby S.J.;
RT "Cloning and sequencing of the gene encoding the Rpos (KatF) sigma
factor from Salmonella typhimurium 14028s.";
RL Biochim. Biophys. Acta 1219:198-200(1994).
RN [3]
RP SEQUENCE OF 298-377 FROM N.A.
RC STRAIN=C52;
RX MEDLINE=95050252; PubMed=7961444;
RA Kowarz L., Coyneault C., Robbe-Saule V., Norel F.;
RT "The Salmonella typhimurium katF (rpos) gene: cloning, nucleotide
sequence, and regulation of spv and spvABC virulence plasmid
genes.";
RL J. Bacteriol. 176:6852-6860(1994).
RN [4]
RP SEQUENCE OF 176:6852-6860(1994).
CC -1- FUNCTION: MAY BE INVOLVED IN STATIONARY-PHASE SURVIVAL.
CC -1- SUBCELLULAR LOCATION: Attached to the inner membrane by a lipid
anchor (Potential).
CC -1- SIMILARITY: BELONGS TO THE E.COLI NLPD / HAEMOPHILUS LPPB FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi-sib.ch/announce/>
CC or send an email to license@ebi-sib.ch).
CC
DR EMBL; AE008833; AAL21805.1; -
DR EMBL; U050111; -; NOT_ANNOTATED_CDS.
DR EMBL; X77752; -; NOT_ANNOTATED_CDS.
DR MEROPS; M37.0PW; -
DR StyGene; SG10513; nlpD.
DR InterPro; IPR002482; LysM.
DR InterPro; IPR002886; Peptidase_M37.
DR Pfam; PF01476; LysM; 1.
DR Pfam; PF01551; Peptidase_M37; 1.
DR SMART; SM00257; LysM; 1.
DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
KW Inner membrane; Lipoprotein; Repeat; Signal; Complete proteome.
FT SIGNAL 1 25
FT CHAIN 26 377
FT LIPOID 26 26
FT DOMAIN 203 250
FT REPEAT 203 209
FT REPEAT 225 231
FT REPEAT 237 243
FT REPEAT 244 250
FT CONFLICT 377 377 R -> P (IN REF. 2).
SQ SEQUENCE 377 AA; 39641 MW; 27C29D77A145ABF0 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 377;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 81 GSTGTSS 87
Db 348 GSTGTSS 354
|||||
|
RESULT 47
RING1_HUMAN STANDARD; PRT; 377 AA.
AC Q06587;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Polycarb complex protein RING1 (RNFI).
GN RING1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93211912; PubMed=7681583;
RA Lovering R., Hanson I.M., Borden K.L.B., Martin S., O'Reilly N.J.,
RA Evan G.I., Rahman D., Pappin D.J.C., Trowsdale J., Freemont P.S.;
RT "Identification and preliminary characterization of a protein motif
related to the zinc finger.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:2112-2116(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Tubby B.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=97342649; PubMed=9199346;
RA Satijn D.P.E., Gunster M.J., van der Viag J., Hamer K.M., Schul W.,
RA Alkema M.J., Saurin A.J., Freemont P.S., van Driel R., Otte A.P.;
RT "RING1 is associated with the polycarb group protein complex and acts
as a transcriptional repressor.";
RL Mol. Cell. Biol. 17:4105-4113(1997).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=99077945; PubMed=9858531;
RA Satijn D.P.E., Otte A.P.;
RT "RING1 interacts with multiple Polycarb-group proteins and displays
tumorigenic activity.";
RL Mol. Cell. Biol. 19:57-68(1999).
RN [5]
RP IDENTIFICATION IN COMPLEX WITH E2F6; TFDPI; MAX; MGA; E2F6; BAT8;
RX CBX3; RNF2; MBLR; L3MBTL2 AND YAF2.
RX MEDLINE=21999559; PubMed=12004135;
RA Ogawa H., Ishiguro K.-I., Gaubatz S., Livingston D.M., Nakatani Y.;
RT "A complex with chromatin modifiers that occupies E2F- and
Myc-responsive genes in G0 cells.";
RL Science 296:1132-1136(2002).
CC -1- FUNCTION: Involved in maintaining the transcriptionally repressive
CC state of genes. Modifies chromatin, rendering it heritably changed
CC in its expressibility. Found in a complex containing histone H3-K9
CC methyltransferase activity.
CC -1- SUBUNIT: Component of the chromatin-associated polycarb complex
CC (PCG). Part of the E2F6-com-1 complex in G0 phase composed of
CC E2F6, MGA, MAX, TFDPI, CBX3, BAT8, E2F6, RING1, RNF2, MBLR,
CC L3MBTL2 and YAF2.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; Z14000; CAA78389.1; -.
DR EMBL; AL031228; CAA20235.1; -.
DR PIR; A47380; A47380.
DR HSSP; P28990; 1CHC.
DR Genew; HGNC:10018; RING1.
DR MIM; 602045; -.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0016564; F:transcriptional repressor activity; IDA.
DR GO; GO:0008270; F:zinc ion binding activity; NAS.
DR GO; GO:0018568; P:chromatin modification; NAS.
DR InterPro; IPR001841; Znf_RING.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00899; ZF_RING_2; 1.
KW Chromatin regulator; Nuclear protein; Transcription regulation;
KW Repressor; Zinc-finger.
FT ZN_FING 19 59 RING-TYPE.
FT DOMAIN 176 231 GLY-RICH.
FT DOMAIN 285 348 GLY-RICH.
FT DOMAIN 172 175 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ SEQUENCE 377 AA; 39145 MW; F406B93593E0CF69 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 377;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
Db 307 GGGAGGG 313
|||||

RESULT 48

DNJ2_STRCO STANDARD; PRT; 378 AA.
AC Q9R0D7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Chapterone protein dnaJ2.
GN DNAJ2 OR SCO2554 OR SCC77.21C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Harris D.E., Quail M.A., Kiese H.,
RA Thomson N.R., James K.D., Carrido-Tarraga A.-M., Challis G.L.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kiese H., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinovitch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Taylor K.,
RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -!- FUNCTION: ACTS AS A CO-CHAPERONE. STIMULATES, JOINTLY WITH GRPE,
CC THE ATPASE ACTIVITY OF DNAK (BY SIMILARITY).
CC -!- COFACTOR: BINDS TWO ZINC IONS PER MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE DNAJ FAMILY.
CC -!- SIMILARITY: Contains 1 J domain.
CC -!- SIMILARITY: Contains 1 CR domain.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; AL939113; CAB66232.1; -.
DR HSSP; P08622; 1XBL.
DR InterPro; IPR002939; DnaJ_C.
DR InterPro; IPR001305; DnaJ_CXXCXGKG.
DR InterPro; IPR001623; DnaJ_N.
DR InterPro; IPR003095; Hsp_DnaJ.
DR Pfam; PF00226; DnaJ_1.
DR Pfam; PF01556; DnaJ_C; 1.
DR Pfam; PF00684; DnaJ_CXXCXGKG; 1.
DR PRINTS; PR00625; DnaJ_PROTEIN.
DR SMART; SM00271; DnaJ; 1.
DR PROSITE; PS00636; DnaJ_1; 1.
DR PROSITE; PS00076; DnaJ_2; 1.
KW Chapterone; DNA replication; Heat shock; Repeat; Zinc; Metal-binding;
KW Complete proteome.
FT DOMAIN 4 68 J-DOMAIN.
FT DOMAIN 75 113 GLY-RICH.
FT REPEAT 143 150 CXXCXGKG MOTIF.
FT REPEAT 160 167 CXXCXGKG MOTIF.
FT REPEAT 186 193 CXXCXGKG MOTIF.
FT REPEAT 200 207 CXXCXGKG MOTIF.
FT METAL 143 143 ZINC 1 (BY SIMILARITY).
FT METAL 146 146 ZINC 1 (BY SIMILARITY).
FT METAL 160 160 ZINC 2 (BY SIMILARITY).
FT METAL 163 163 ZINC 2 (BY SIMILARITY).
FT METAL 186 186 ZINC 2 (BY SIMILARITY).
FT METAL 189 189 ZINC 2 (BY SIMILARITY).
FT METAL 200 200 ZINC 1 (BY SIMILARITY).
FT METAL 203 203 ZINC 1 (BY SIMILARITY).
SQ SEQUENCE 378 AA; 40488 MW; 8C3DEF5B4CC6808D CRC64;

Query Match 1.9%; Score 7; DB 1; Length 378;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGG 240
Db 75 AGGGAGG 81
|||||

RESULT 49

NLPD_ECOLI STANDARD; PRT; 379 AA.
AC P33648;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Lipoprotein nlpD precursor.
GN NLPD OR B2742 OR SF2765.
OS Escherichia coli and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 623;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=MP180;
RX MEDLINE=94179096; PubMed=8132457;
RA Ichikawa J.K., Li C., Fu J.C., Clarke S.;
RT "A gene at 59 minutes on the Escherichia coli chromosome encodes a
RT lipoprotein with unusual amino acid repeat sequences.";
RL J. Bacteriol. 176:1630-1638(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;

RA MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.P.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of *Escherichia coli* K-12.";
RL Science 277:1453-1474(1997).
[3]
RN SEQUENCE OF 99-379 FROM N.A.
RN SPECIES=E.coli; STRAIN=K12 / DH1;
RC MEDLINE=94268497; PubMed=8208244;
RA Takayanagi Y., Tanaka K., Takahashi H.;
RT "Structure of the 5' upstream region and the regulation of the rpos
RL gene of *Escherichia coli*.";
RL Mol. Gen. Genet. 243:525-531(1994).
[4]
RN SEQUENCE FROM N.A.
RN SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RC MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity
RT through comparison with genomes of *Escherichia coli* K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
CC -!- FUNCTION: MAY BE INVOLVED IN STATIONARY-PHASE SURVIVAL.
CC -!- SUBCELLULAR LOCATION: Attached to the inner membrane by a lipid
CC anchor (Potential).
CC -!- SIMILARITY: BELONGS TO THE E-COLI NLPD / HAEMOPHILUS LPPB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L07869; AAA17875.1; -;
DR EMBL; U29579; AAA69252.1; -;
DR EMBL; AE000358; AAC75784.1; -;
DR EMBL; D17549; BAA04487.1; -;
DR EMBL; AE015290; AAN44254.1; ALT_INIT.
DR PIR; B55522; B55522.
DR MEROPS; M37.UFW; -;
DR SWISS-2DPAGE; P33648; COLI.
DR EcoGene; EG12111; nlpD.
DR InterPro; IPR002482; LysM.
DR InterPro; IPR002886; Peptidase_M37.
DR Pfam; PF01476; LysM.
DR Pfam; PF01551; Peptidase_M37; 1.
DR SMART; SM00257; LysM; 1.
DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
KW Inner membrane; Lipoprotein; Repeat; Signal; Complete proteome.
FT SIGNAL 1 25
FT CHAIN 26 379 LIPOPROTEIN NLPD.
FT LIPID 26 26 N-ACYL DIGLYCERIDE.
FT DOMAIN 66 97 4 X 8 AA TANDEM REPEATS OF Q-Q-P-Q-I-Q-
FT P-V.
FT REPEAT 66 73 1-1.
FT REPEAT 74 81 1-2 (APPROXIMATE).
FT REPEAT 82 89 1-3.
FT REPEAT 90 97 1-4 (APPROXIMATE).
FT DOMAIN 205 252 4 X 7 AA APPROXIMATE REPEATS.
FT REPEAT 205 211 2-1.
FT REPEAT 227 233 2-2.
FT REPEAT 239 245 2-3.
FT REPEAT 246 252 2-4.
FT CONFLICT 139 139 G -> A (IN REF. 3).
SQ SEQUENCE 379 AA; 40149 MW; A8B6A2B8456105FE CRC64;

Query Match 1.9%; Score 7; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 81 GSTGTSS 87
DB 350 GSTGTSS 356
RESULT 50
TCF7_HUMAN STANDARD; PRT; 383 AA.
ID TCF7_HUMAN
AC P36402; Q9UKI4;
DT 01-JUN-1994 (Rel. 29, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Transcription factor 7 (T-cell-specific transcription factor 1) (TCF-
DE 1) (T-cell factor 1).
GN TCF7 OR TCF1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 4L).
RX MEDLINE=91114695; PubMed=1989880;
RA van de Wetering M., Oosterwegel M.A., Dooijes D., Clevers H.;
RT "Identification and cloning of TCF-1, a T lymphocyte-specific
RT transcription factor containing a sequence-specific HMG box.";
RL EMBO J. 10:123-132(1991).
RN [2]
RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
RX MEDLINE=92235082; PubMed=1569101;
RA van de Wetering M., Oosterwegel M.A., Holstege F., Dooyes D.,
RA Suifkerbuijk R., Geurts van Kessel A., Clevers H.;
RT "The human T cell transcription factor-1 gene. Structure,
RT localization, and promoter characterization.";
RL J. Biol. Chem. 267:8530-8536(1992).
RN [3]
RP SEQUENCE OF 1-68 FROM N.A. (ISOFORM 4L), TISSUE SPECIFICITY, AND
RP INDUCTION BY TCF7L2 AND CTNNB1.
RX MEDLINE=99420370; PubMed=10489374;
RA Roose J., Huls G., van Beest M., Moerer P., van der Horn K.,
RA Goldschmeding R., Logtenberg T., Clevers H.;
RT "Synergy between tumor suppressor APC and the beta-catenin-Tcf4 target
RT Tcf1.";
RL Science 285:1923-1926(1999).
RN [4]
RP UTILISATION OF AN UPSTREAM PROMOTER, AND IDENTIFICATION OF ISOFORMS
RP 1S AND 4L.
RX MEDLINE=96182076; PubMed=8622675;
RA van de Wetering M., Castrop J., Korinek V., Clevers H.;
RT "Extensive alternative splicing and dual promoter usage generate Tcf-1
RT protein isoforms with differential transcription control properties.";
RL Mol. Cell. Biol. 16:745-752(1996).
RN [5]
RP INTERACTION WITH CTNNB1.
RX MEDLINE=98147763; PubMed=9488439;
RA Korinek V., Barker N., Willert K., Molenaar M., Roose J., Wagenaar G.,
RA Markman M., Lammers W., Destree O., Clevers H.;
RT "Two members of the Tcf family implicated in Wnt/b-catenin signaling
RT during embryogenesis in the mouse.";
RL Mol. Cell. Biol. 18:1248-1256(1998).
RN [6]
RP INTERACTION WITH CTNNB1 AND AES.
RX MEDLINE=98454908; PubMed=9783587;
RA Roose J., Molenaar M., Peterson J., Hurenkamp J., Brantjes H.,
RA Moerer P., van de Wetering M., Destree O., Clevers H.;
RT "The Xenopus Wnt effector XTcf-3 interacts with Groucho-related
RT transcriptional repressors.";
RL Nature 395:608-612(1998).
RN [7]
RP INTERACTION WITH AES; TLE1; TLE2; TLE3 AND TLE4.
RN

```

RX MEDLINE=21169341; PubMed=11266540;
RA Brantjes H., Roese J., van De Wetering M., Clevers H.;
RT "All Tcf HMG box transcription factors interact with Groucho-related
RL co-repressors."
RL Nucleic Acids Res. 29:1410-1419(2001).
CC
CC -!- FUNCTION: Transcriptional activator involved in T-cell lymphocyte
CC differentiation. Necessary for the survival of CD4(+) CD8(+)
CC immature thymocytes. Isoforms lacking the N-terminal CTNNB1
CC binding domain cannot fulfill this role. Binds to the T-
CC lymphocyte-specific enhancer element (5-WCAGAAG-3') found in the
CC promoter of the CD3E gene. May also act as feedback
CC transcriptional repressor of CTNNB1 and TCF7L2 target genes (By
CC similarity). TLE1, TLE2, TLE3 and TLE4 repress transactivation
CC mediated by TCF7 and CTNNB1.
CC
CC -!- SUBUNIT: Binds the armadillo repeat of CTNNB1 and forms a stable
CC complex. Interacts with AEB, TLE1, TLE2, TLE3 and TLE4.
CC
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC
CC -!- ALTERNATIVE PRODUCTS:
CC
CC Event=Alternative promoter;
CC Comment=2 series of isoforms, L (shown here) and S, are
CC produced by use of alternative promoters;
CC Event=Alternative splicing; Named isoforms=8;
CC Comment=Additional isoforms seem to exist;
CC
CC Name=4L; Synonyms=B;
CC IsoId=P36402-1; Sequence=Displayed;
CC Name=4S;
CC IsoId=P36402-2; Sequence=VSP_006961;
CC Name=1L;
CC IsoId=P36402-3; Sequence=VSP_006960;
CC Note=Produced by alternative splicing of isoform 4L;
CC Name=1S; Synonyms=B;
CC IsoId=P36402-4; Sequence=VSP_006961, VSP_006960;
CC Note=Produced by alternative splicing of isoform 4S;
CC Name=2L;
CC IsoId=P36402-5; Sequence=VSP_002191;
CC Note=Produced by alternative splicing of isoform 4L;
CC Name=2S; Synonyms=B;
CC IsoId=P36402-6; Sequence=VSP_006961, VSP_002191;
CC Note=Produced by alternative splicing of isoform 4S;
CC Name=3L;
CC IsoId=P36402-7; Sequence=VSP_002192;
CC Note=Produced by alternative splicing of isoform 4L;
CC Name=3S; Synonyms=C;
CC IsoId=P36402-8; Sequence=VSP_006961, VSP_002192;
CC Note=Produced by alternative splicing of isoform 4S;
CC -!- TISSUE SPECIFICITY: Predominantly in T-cells. Also detected in
CC proliferating intestinal epithelial cells and in the basal
CC epithelial cells of mammary gland epithelium.
CC
CC -!- INDUCTION: By TCF7L2 and CTNNB1.
CC
CC -!- SIMILARITY: Belongs to the TCF/Lef family.
CC
CC -!- SIMILARITY: Contains 1 HMG box domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/annouce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X59869; CAA42526.1; -
CC EMBL; X59870; CAA42527.1; -
CC EMBL; X59871; CAA42528.1; -
CC EMBL; X63901; CAA56795.1; -
CC EMBL; X63902; CAA56796.1; -
CC EMBL; X63903; CAA56797.1; -
CC EMBL; X63904; CAA56798.1; -
CC EMBL; X63905; CAA56799.1; -
CC EMBL; X63906; CAA56800.1; -
CC EMBL; X63907; CAA56801.1; -
CC EMBL; X63908; CAA56802.1; -
CC EMBL; X63909; CAA56803.1; -
CC EMBL; X63910; CAA56804.1; -
CC EMBL; X63911; CAA56805.1; -
CC EMBL; X63912; CAA56806.1; -
CC EMBL; X63913; CAA56807.1; -
CC EMBL; X63914; CAA56808.1; -
CC EMBL; X63915; CAA56809.1; -
CC EMBL; X63916; CAA56810.1; -
CC EMBL; X63917; CAA56811.1; -
CC EMBL; X63918; CAA56812.1; -
CC EMBL; X63919; CAA56813.1; -
CC EMBL; X63920; CAA56814.1; -
CC EMBL; X63921; CAA56815.1; -
CC EMBL; X63922; CAA56816.1; -
CC EMBL; X63923; CAA56817.1; -
CC EMBL; X63924; CAA56818.1; -
CC EMBL; X63925; CAA56819.1; -
CC EMBL; X63926; CAA56820.1; -
CC EMBL; X63927; CAA56821.1; -
CC EMBL; X63928; CAA56822.1; -
CC EMBL; X63929; CAA56823.1; -
CC EMBL; X63930; CAA56824.1; -
CC EMBL; X63931; CAA56825.1; -
CC EMBL; X63932; CAA56826.1; -
CC EMBL; X63933; CAA56827.1; -
CC EMBL; X63934; CAA56828.1; -
CC EMBL; X63935; CAA56829.1; -
CC EMBL; X63936; CAA56830.1; -
CC EMBL; X63937; CAA56831.1; -
CC EMBL; X63938; CAA56832.1; -
CC EMBL; X63939; CAA56833.1; -
CC EMBL; X63940; CAA56834.1; -
CC EMBL; X63941; CAA56835.1; -
CC EMBL; X63942; CAA56836.1; -
CC EMBL; X63943; CAA56837.1; -
CC EMBL; X63944; CAA56838.1; -
CC EMBL; X63945; CAA56839.1; -
CC EMBL; X63946; CAA56840.1; -
CC EMBL; X63947; CAA56841.1; -
CC EMBL; X63948; CAA56842.1; -
CC EMBL; X63949; CAA56843.1; -
CC EMBL; X63950; CAA56844.1; -
CC EMBL; X63951; CAA56845.1; -
CC EMBL; X63952; CAA56846.1; -
CC EMBL; X63953; CAA56847.1; -
CC EMBL; X63954; CAA56848.1; -
CC EMBL; X63955; CAA56849.1; -
CC EMBL; X63956; CAA56850.1; -
CC EMBL; X63957; CAA56851.1; -
CC EMBL; X63958; CAA56852.1; -
CC EMBL; X63959; CAA56853.1; -
CC EMBL; X63960; CAA56854.1; -
CC EMBL; X63961; CAA56855.1; -
CC EMBL; X63962; CAA56856.1; -
CC EMBL; X63963; CAA56857.1; -
CC EMBL; X63964; CAA56858.1; -
CC EMBL; X63965; CAA56859.1; -
CC EMBL; X63966; CAA56860.1; -
CC EMBL; X63967; CAA56861.1; -
CC EMBL; X63968; CAA56862.1; -
CC EMBL; X63969; CAA56863.1; -
CC EMBL; X63970; CAA56864.1; -
CC EMBL; X63971; CAA56865.1; -
CC EMBL; X63972; CAA56866.1; -
CC EMBL; X63973; CAA56867.1; -
CC EMBL; X63974; CAA56868.1; -
CC EMBL; X63975; CAA56869.1; -
CC EMBL; X63976; CAA56870.1; -
CC EMBL; X63977; CAA56871.1; -
CC EMBL; X63978; CAA56872.1; -
CC EMBL; X63979; CAA56873.1; -
CC EMBL; X63980; CAA56874.1; -
CC EMBL; X63981; CAA56875.1; -
CC EMBL; X63982; CAA56876.1; -
CC EMBL; X63983; CAA56877.1; -
CC EMBL; X63984; CAA56878.1; -
CC EMBL; X63985; CAA56879.1; -
CC EMBL; X63986; CAA56880.1; -
CC EMBL; X63987; CAA56881.1; -
CC EMBL; X63988; CAA56882.1; -
CC EMBL; X63989; CAA56883.1; -
CC EMBL; X63990; CAA56884.1; -
CC EMBL; X63991; CAA56885.1; -
CC EMBL; X63992; CAA56886.1; -
CC EMBL; X63993; CAA56887.1; -
CC EMBL; X63994; CAA56888.1; -
CC EMBL; X63995; CAA56889.1; -
CC EMBL; X63996; CAA56890.1; -
CC EMBL; X63997; CAA56891.1; -
CC EMBL; X63998; CAA56892.1; -
CC EMBL; X63999; CAA56893.1; -
CC EMBL; X64000; CAA56894.1; -
CC EMBL; X64001; CAA56895.1; -
CC EMBL; X64002; CAA56896.1; -
CC EMBL; X64003; CAA56897.1; -
CC EMBL; X64004; CAA56898.1; -
CC EMBL; X64005; CAA56899.1; -
CC EMBL; X64006; CAA56900.1; -
CC EMBL; X64007; CAA56901.1; -
CC EMBL; X64008; CAA56902.1; -
CC EMBL; X64009; CAA56903.1; -
CC EMBL; X64010; CAA56904.1; -
CC EMBL; X64011; CAA56905.1; -
CC EMBL; X64012; CAA56906.1; -
CC EMBL; X64013; CAA56907.1; -
CC EMBL; X64014; CAA56908.1; -
CC EMBL; X64015; CAA56909.1; -
CC EMBL; X64016; CAA56910.1; -
CC EMBL; X64017; CAA56911.1; -
CC EMBL; X64018; CAA56912.1; -
CC EMBL; X64019; CAA56913.1; -
CC EMBL; X64020; CAA56914.1; -
CC EMBL; X64021; CAA56915.1; -
CC EMBL; X64022; CAA56916.1; -
CC EMBL; X64023; CAA56917.1; -
CC EMBL; X64024; CAA56918.1; -
CC EMBL; X64025; CAA56919.1; -
CC EMBL; X64026; CAA56920.1; -
CC EMBL; X64027; CAA56921.1; -
CC EMBL; X64028; CAA56922.1; -
CC EMBL; X64029; CAA56923.1; -
CC EMBL; X64030; CAA56924.1; -
CC EMBL; X64031; CAA56925.1; -
CC EMBL; X64032; CAA56926.1; -
CC EMBL; X64033; CAA56927.1; -
CC EMBL; X64034; CAA56928.1; -
CC EMBL; X64035; CAA56929.1; -
CC EMBL; X64036; CAA56930.1; -
CC EMBL; X64037; CAA56931.1; -
CC EMBL; X64038; CAA56932.1; -
CC EMBL; X64039; CAA56933.1; -
CC EMBL; X64040; CAA56934.1; -
CC EMBL; X64041; CAA56935.1; -
CC EMBL; X64042; CAA56936.1; -
CC EMBL; X64043; CAA56937.1; -
CC EMBL; X64044; CAA56938.1; -
CC EMBL; X64045; CAA56939.1; -
CC EMBL; X64046; CAA56940.1; -
CC EMBL; X64047; CAA56941.1; -
CC EMBL; X64048; CAA56942.1; -
CC EMBL; X64049; CAA56943.1; -
CC EMBL; X64050; CAA56944.1; -
CC EMBL; X64051; CAA56945.1; -
CC EMBL; X64052; CAA56946.1; -
CC EMBL; X64053; CAA56947.1; -
CC EMBL; X64054; CAA56948.1; -
CC EMBL; X64055; CAA56949.1; -
CC EMBL; X64056; CAA56950.1; -
CC EMBL; X64057; CAA56951.1; -
CC EMBL; X64058; CAA56952.1; -
CC EMBL; X64059; CAA56953.1; -
CC EMBL; X64060; CAA56954.1; -
CC EMBL; X64061; CAA56955.1; -
CC EMBL; X64062; CAA56956.1; -
CC EMBL; X64063; CAA56957.1; -
CC EMBL; X64064; CAA56958.1; -
CC EMBL; X64065; CAA56959.1; -
CC EMBL; X64066; CAA56960.1; -
CC EMBL; X64067; CAA56961.1; -
CC EMBL; X64068; CAA56962.1; -
CC EMBL; X64069; CAA56963.1; -
CC EMBL; X64070; CAA56964.1; -
CC EMBL; X64071; CAA56965.1; -
CC EMBL; X64072; CAA56966.1; -
CC EMBL; X64073; CAA56967.1; -
CC EMBL; X64074; CAA56968.1; -
CC EMBL; X64075; CAA56969.1; -
CC EMBL; X64076; CAA56970.1; -
CC EMBL; X64077; CAA56971.1; -
CC EMBL; X64078; CAA56972.1; -
CC EMBL; X64079; CAA56973.1; -
CC EMBL; X64080; CAA56974.1; -
CC EMBL; X64081; CAA56975.1; -
CC EMBL; X64082; CAA56976.1; -
CC EMBL; X64083; CAA56977.1; -
CC EMBL; X64084; CAA56978.1; -
CC EMBL; X64085; CAA56979.1; -
CC EMBL; X64086; CAA56980.1; -
CC EMBL; X64087; CAA56981.1; -
CC EMBL; X64088; CAA56982.1; -
CC EMBL; X64089; CAA56983.1; -
CC EMBL; X64090; CAA56984.1; -
CC EMBL; X64091; CAA56985.1; -
CC EMBL; X64092; CAA56986.1; -
CC EMBL; X64093; CAA56987.1; -
CC EMBL; X64094; CAA56988.1; -
CC EMBL; X64095; CAA56989.1; -
CC EMBL; X64096; CAA56990.1; -
CC EMBL; X64097; CAA56991.1; -
CC EMBL; X64098; CAA56992.1; -
CC EMBL; X64099; CAA56993.1; -
CC EMBL; X64100; CAA56994.1; -
CC EMBL; X64101; CAA56995.1; -
CC EMBL; X64102; CAA56996.1; -
CC EMBL; X64103; CAA56997.1; -
CC EMBL; X64104; CAA56998.1; -
CC EMBL; X64105; CAA56999.1; -
CC EMBL; X64106; CAA57000.1; -
CC EMBL; X64107; CAA57001.1; -
CC EMBL; X64108; CAA57002.1; -
CC EMBL; X64109; CAA57003.1; -
CC EMBL; X64110; CAA57004.1; -
CC EMBL; X64111; CAA57005.1; -
CC EMBL; X64112; CAA57006.1; -
CC EMBL; X64113; CAA57007.1; -
CC EMBL; X64114; CAA57008.1; -
CC EMBL; X64115; CAA57009.1; -
CC EMBL; X64116; CAA57010.1; -
CC EMBL; X64117; CAA57011.1; -
CC EMBL; X64118; CAA57012.1; -
CC EMBL; X64119; CAA57013.1; -
CC EMBL; X64120; CAA57014.1; -
CC EMBL; X64121; CAA57015.1; -
CC EMBL; X64122; CAA57016.1; -
CC EMBL; X64123; CAA57017.1; -
CC EMBL; X64124; CAA57018.1; -
CC EMBL; X64125; CAA57019.1; -
CC EMBL; X64126; CAA57020.1; -
CC EMBL; X64127; CAA57021.1; -
CC EMBL; X64128; CAA57022.1; -
CC EMBL; X64129; CAA57023.1; -
CC EMBL; X64130; CAA57024.1; -
CC EMBL; X64131; CAA57025.1; -
CC EMBL; X64132; CAA57026.1; -
CC EMBL; X64133; CAA57027.1; -
CC EMBL; X64134; CAA57028.1; -
CC EMBL; X64135; CAA57029.1; -
CC EMBL; X64136; CAA57030.1; -
CC EMBL; X64137; CAA57031.1; -
CC EMBL; X64138; CAA57032.1; -
CC EMBL; X64139; CAA57033.1; -
CC EMBL; X64140; CAA57034.1; -
CC EMBL; X64141; CAA57035.1; -
CC EMBL; X64142; CAA57036.1; -
CC EMBL; X64143; CAA57037.1; -
CC EMBL; X64144; CAA57038.1; -
CC EMBL; X64145; CAA57039.1; -
CC EMBL; X64146; CAA57040.1; -
CC EMBL; X64147; CAA57041.1; -
CC EMBL; X64148; CAA57042.1; -
CC EMBL; X64149; CAA57043.1; -
CC EMBL; X64150; CAA57044.1; -
CC EMBL; X64151; CAA57045.1; -
CC EMBL; X64152; CAA57046.1; -
CC EMBL; X64153; CAA57047.1; -
CC EMBL; X64154; CAA57048.1; -
CC EMBL; X64155; CAA57049.1; -
CC EMBL; X64156; CAA57050.1; -
CC EMBL; X64157; CAA57051.1; -
CC EMBL; X64158; CAA57052.1; -
CC EMBL; X64159; CAA57053.1; -
CC EMBL; X64160; CAA57054.1; -
CC EMBL; X64161; CAA57055.1; -
CC EMBL; X64162; CAA57056.1; -
CC EMBL; X64163; CAA57057.1; -
CC EMBL; X64164; CAA57058.1; -
CC EMBL; X64165; CAA57059.1; -
CC EMBL; X64166; CAA57060.1; -
CC EMBL; X64167; CAA57061.1; -
CC EMBL; X64168; CAA57062.1; -
CC EMBL; X64169; CAA57063.1; -
CC EMBL; X64170; CAA57064.1; -
CC EMBL; X64171; CAA57065.1; -
CC EMBL; X64172; CAA57066.1; -
CC EMBL; X64173; CAA57067.1; -
CC EMBL; X64174; CAA57068.1; -
CC EMBL; X64175; CAA57069.1; -
CC EMBL; X64176; CAA57070.1; -
CC EMBL; X64177; CAA57071.1; -
CC EMBL; X64178; CAA57072.1; -
CC EMBL; X64179; CAA57073.1; -
CC EMBL; X64180; CAA57074.1; -
CC EMBL; X64181; CAA57075.1; -
CC EMBL; X64182; CAA57076.1; -
CC EMBL; X64183; CAA57077.1; -
CC EMBL; X64184; CAA57078.1; -
CC EMBL; X64185; CAA57079.1; -
CC EMBL; X64186; CAA57080.1; -
CC EMBL; X64187; CAA57081.1; -
CC EMBL; X64188; CAA57082.1; -
CC EMBL; X64189; CAA57083.1; -
CC EMBL; X64190; CAA57084.1; -
CC EMBL; X64191; CAA57085.1; -
CC EMBL; X64192; CAA57086.1; -
CC EMBL; X64193; CAA57087.1; -
CC EMBL; X64194; CAA57088.1; -
CC EMBL; X64195; CAA57089.1; -
CC EMBL; X64196; CAA57090.1; -
CC EMBL; X64197; CAA57091.1; -
CC EMBL; X64198; CAA57092.1; -
CC EMBL; X64199; CAA57093.1; -
CC EMBL; X64200; CAA57094.1; -
CC EMBL; X64201; CAA57095.1; -
CC EMBL; X64202; CAA57096.1; -
CC EMBL; X64203; CAA57097.1; -
CC EMBL; X64204; CAA57098.1; -
CC EMBL; X64205; CAA57099.1; -
CC EMBL; X64206; CAA57100.1; -
CC EMBL; X64207; CAA57101.1; -
CC EMBL; X64208; CAA57102.1; -
CC EMBL; X64209; CAA57103.1; -
CC EMBL; X64210; CAA57104.1; -
CC EMBL; X64211; CAA57105.1; -
CC EMBL; X64212; CAA57106.1; -
CC EMBL; X64213; CAA57107.1; -
CC EMBL; X64214; CAA57108.1; -
CC EMBL; X64215; CAA57109.1; -
CC EMBL; X64216; CAA57110.1; -
CC EMBL; X64217; CAA57111.1; -
CC EMBL; X64218; CAA57112.1; -
CC EMBL; X64219; CAA57113.1; -
CC EMBL; X64220; CAA57114.1; -
CC EMBL; X64221; CAA57115.1; -
CC EMBL; X64222; CAA57116.1; -
CC EMBL; X64223; CAA57117.1; -
CC EMBL; X64224; CAA57118.1; -
CC EMBL; X64225; CAA57119.1; -
CC EMBL; X64226; CAA57120.1; -
CC EMBL; X64227; CAA57121.1; -
CC EMBL; X64228; CAA57122.1; -
CC EMBL; X64229; CAA57123.1; -
CC EMBL; X64230; CAA57124.1; -
CC EMBL; X64231; CAA57125.1; -
CC EMBL; X64232; CAA57126.1; -
CC EMBL; X64233; CAA57127.1; -
CC EMBL; X64234; CAA57128.1; -
CC EMBL; X64235; CAA57129.1; -
CC EMBL; X64236; CAA57130.1; -
CC EMBL; X64237; CAA57131.1; -
CC EMBL; X64238; CAA57132.1; -
CC EMBL; X64239; CAA57133.1; -
CC EMBL; X64240; CAA57134.1; -
CC EMBL; X64241; CAA57135.1; -
CC EMBL; X64242; CAA57136.1; -
CC EMBL; X64243; CAA57137.1; -
CC EMBL; X64244; CAA57138.1; -
CC EMBL; X64245; CAA57139.1; -
CC EMBL; X64246; CAA57140.1; -
CC EMBL; X64247; CAA57141.1; -
CC EMBL; X64248; CAA57142.1; -
CC EMBL; X64249; CAA57143.1; -
CC EMBL; X64250; CAA57144.1; -
CC EMBL; X64251; CAA57145.1; -
CC EMBL; X64252; CAA57146.1; -
CC EMBL; X64253; CAA57147.1; -
CC EMBL; X64254; CAA57148.1; -
CC EMBL; X64255; CAA57149.1; -
CC EMBL; X64256; CAA57150.1; -
CC EMBL; X64257; CAA57151.1; -
CC EMBL; X64258; CAA57152.1; -
CC EMBL; X64259; CAA57153.1; -
CC EMBL; X64260; CAA57154.1; -
CC EMBL; X64261; CAA57155.1; -
CC EMBL; X64262; CAA57156.1; -
CC EMBL; X64263; CAA57157.1; -
CC EMBL; X64264; CAA57158.1; -
CC EMBL; X64265; CAA57159.1; -
CC EMBL; X64266; CAA57160.1; -
CC EMBL; X64267; CAA57161.1; -
CC EMBL; X64268; CAA57162.1; -
CC EMBL; X64269; CAA57163.1; -
CC EMBL; X64270; CAA57164.1; -
CC EMBL; X64271; CAA57165.1; -
CC EMBL; X64272; CAA57166.1; -
CC EMBL; X64273; CAA57167.1; -
CC EMBL; X64274; CAA57168.1; -
CC EMBL; X64275; CAA57169.1; -
CC EMBL; X64276; CAA57170.1; -
CC EMBL; X64277; CAA57171.1; -
CC EMBL; X64278; CAA57172.1; -
CC EMBL; X64279; CAA57173.1; -
CC EMBL; X64280; CAA57174.1; -
CC EMBL; X64281; CAA57175.1; -
CC EMBL; X64282; CAA57176.1; -
CC EMBL; X64283; CAA57177.1; -
CC EMBL; X64284; CAA57178.1; -
CC EMBL; X64285; CAA57179.1; -
CC EMBL; X64286; CAA57180.1; -
CC EMBL; X64287; CAA57181.1; -
CC EMBL; X64288; CAA57182.1; -
CC EMBL; X64289; CAA57183.1; -
CC EMBL; X64290; CAA57184.1; -
CC EMBL; X64291; CAA57185.1; -
CC EMBL; X64292; CAA57186.1; -
CC EMBL; X64293; CAA57187.1; -
CC EMBL; X64294; CAA57188.1; -
CC EMBL; X64295; CAA57189.1; -
CC EMBL; X64296; CAA57190.1; -
CC EMBL; X64297; CAA57191.1; -
CC EMBL; X64298; CAA57192.1; -
CC EMBL; X64299; CAA57193.1; -
CC EMBL; X64300; CAA57194.1; -
CC EMBL; X64301; CAA57195.1; -
CC EMBL; X64302; CAA57196.1; -
CC EMBL; X64303; CAA57197.1; -
CC EMBL; X64304; CAA57198.1; -
CC EMBL; X64305; CAA57199.1; -
CC EMBL; X64306; CAA57200.1; -
CC EMBL; X64307; CAA57201.1; -
CC EMBL; X64308; CAA57202.1; -
CC EMBL; X64309; CAA57203.1; -
CC EMBL; X64310; CAA57204.1; -
CC EMBL; X64311; CAA57205.1; -
CC EMBL; X64312; CAA57206.1; -
CC EMBL; X64313; CAA57207.1; -
CC EMBL; X64314; CAA57208.1; -
CC EMBL; X64315; CAA57209.1; -
CC EMBL; X64316; CAA57210.1; -
CC EMBL; X64317; CAA57211.1; -
CC EMBL; X64318; CAA57212.1; -
CC EMBL; X64319; CAA57213.1; -
CC EMBL; X64320; CAA57214.1; -
CC EMBL; X64321; CAA57215.1; -
CC EMBL; X64322; CAA57216.1; -
CC EMBL; X64323; CAA57217.1; -
CC EMBL; X64324; CAA57218.1; -
CC EMBL; X64325; CAA57219.1; -
CC EMBL; X64326; CAA57220.1; -
CC EMBL; X64327; CAA57221.1; -
CC EMBL; X64328; CAA57222.1; -
CC EMBL; X64329; CAA57223.1; -
CC EMBL; X64330; CAA57224.1; -
CC EMBL; X64331; CAA57225.1; -
CC EMBL; X64332; CAA57226.1; -
CC EMBL; X64333; CAA57227.1; -
CC EMBL; X64334; CAA57228.1; -
CC EMBL; X64335; CAA57229.1; -
CC EMBL; X64336; CAA57230.1; -
CC EMBL; X64337; CAA57231.1; -
CC EMBL; X64338; CAA57232.1; -
CC EMBL; X64339; CAA57233.1; -
CC EMBL; X64340; CAA57234.1; -
CC EMBL; X64341; CAA57235.1; -
CC EMBL; X64342; CAA57236.1; -
CC EMBL; X64343; CAA57237.1; -
CC EMBL; X64344; CAA57238.1; -
CC EMBL; X64345; CAA57239.1; -
CC EMBL; X64346; CAA57240.1; -
CC EMBL; X64347; CAA57241.1; -
CC EMBL; X64348; CAA57242.1; -
CC EMBL; X64349; CAA57243.1; -
CC EMBL; X64350; CAA57244.1; -
CC EMBL; X64351; CAA57245.1; -
CC EMBL; X64352; CAA57246.1; -
CC EMBL; X64353; CAA57247.1; -
CC EMBL; X64354; CAA57248.1; -
CC EMBL; X64355; CAA57249.1; -
CC EMBL; X64356; CAA57250.1; -
CC EMBL; X64357; CAA57251.1; -
CC EMBL; X64358; CAA57252.1; -
CC EMBL; X64359; CAA57253.1; -
CC EMBL; X64360; CAA57254.1; -
CC EMBL; X64361; CAA57255.1; -
CC EMBL; X64362; CAA57256.1
```


use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; L02957; AAA49950.1; --
PIR; S40778; S40778.
HSSP; P09651; IUP1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS0102; RRM; 2.
DR PROSITE; PS00030; RRM RNP 1; 2.
KW Nuclear protein; RNA-binding; Ribonucleoprotein; Repeat.
FT DOMAIN 27 110 RNA-BINDING (RRM) 1.
FT DOMAIN 118 206 RNA-BINDING (RRM) 2.
FT DOMAIN 207 385 GLY-RICH.
SQ SEQUENCE 385 AA; 39438 MW; 96895F1B997C9E9D CRC64;

Query Match 1.9%; Score 7; DB 1; Length 385;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
DB 302 GGGAGGG 308
|||||

RESULT 52
ID_DNAJ_HALCU STANDARD; PRT; 389 AA.
AC O34135;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chapterone protein dnaJ (Heat shock protein 40).
GN DnaJ OR Hsp40.
OS Halobacterium cutirubrum.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 33170;
RX MEDLINE=97383250; PubMed=9236279;
RA Bustard K., Gupta R.S.;
RT "The sequences of heat shock protein 40 (DnaJ) homologs provide evidence for a close evolutionary relationship between the Deinococcus-thermus group and cyanobacteria.";
RL J. Mol. Evol. 45:193-205(1997).
CC -!- FUNCTION: ACTS AS A CO-CHAPERONE. STIMULATES, JOINTLY WITH GRPE, THE ATPASE ACTIVITY OF DnaK (BY SIMILARITY).
CC -!- COFACTOR: BINDS TWO ZINC IONS PER MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE DnaJ FAMILY.
CC -!- SIMILARITY: Contains 1 J domain.
CC -!- SIMILARITY: Contains 1 CR domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; U93357; AAB96891.1; --
PIR; T44957; T44957.
HSSP; P08622; 1BQ2.
DR InterPro; IPR002939; DnaJ C.
DR InterPro; IPR001305; DnaJ_CXXCXGKG.
DR InterPro; IPR001623; DnaJ_N.
DR InterPro; IPR003095; Hsp_DnaJ.

Pfam; PF00226; DnaJ; 1.
DR Pfam; PF01556; DnaJ C; 1.
DR Pfam; PF00684; DnaJ_CXXCXGKG; 1.
DR PRINTS; PR00625; DnaJPROTEIN.
DR SMART; SM00271; DnaJ; 1.
DR PROSITE; PS00636; DnaJ 1; FALSE_NEG.
DR PROSITE; PS00076; DnaJ 2; 1.
DR PROSITE; PS00637; DnaJ_CXXCXGKG; FALSE_NEG.
KW Chapterone; DNA replication; Heat shock; Repeat; Zinc; Metal-binding.
FT DOMAIN 4 68 J-DOMAIN.
FT DOMAIN 79 131 GLY-RICH.
FT REPEAT 165 172 CXXCXGKG MOTIF.
FT REPEAT 182 189 CXXCXGKG MOTIF.
FT REPEAT 208 215 CXXCXGKG MOTIF.
FT REPEAT 222 229 CXXCXGKG MOTIF.
FT METAL 165 165 ZINC 1 (BY SIMILARITY).
FT METAL 168 168 ZINC 1 (BY SIMILARITY).
FT METAL 182 182 ZINC 2 (BY SIMILARITY).
FT METAL 185 185 ZINC 2 (BY SIMILARITY).
FT METAL 208 208 ZINC 2 (BY SIMILARITY).
FT METAL 211 211 ZINC 2 (BY SIMILARITY).
FT METAL 222 222 ZINC 1 (BY SIMILARITY).
FT METAL 225 225 ZINC 1 (BY SIMILARITY).
SQ SEQUENCE 389 AA; 41390 MW; 12C1426CBF309633 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 389;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
DB 103 GGGAGGG 109
|||||

RESULT 53
ID_DNAJ_HALN1 STANDARD; PRT; 391 AA.
AC Q9HRV3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chapterone protein dnaJ.
GN DnaJ OR VNG0489G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M., Shukla H.D., LaSkey S.R., Balliga N.S., Thorsson V., Sbrogna J., Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A., Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W., Madlocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H., Isenbarger T.A., Peck R.F., Pohlshocher M., Spudis J.L., Jung K.-H., Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D., Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -!- FUNCTION: ACTS AS A CO-CHAPERONE. STIMULATES, JOINTLY WITH GRPE, THE ATPASE ACTIVITY OF DnaK (BY SIMILARITY).
CC -!- COFACTOR: BINDS TWO ZINC IONS PER MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE DnaJ FAMILY.
CC -!- SIMILARITY: Contains 1 J domain.
CC -!- SIMILARITY: Contains 1 CR domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC -----
CC ENBL; AE005002; AAG19025.1; --
CC PIR; E84207; E84207.
CC HSPP; P08622; BQZ.
CC InterPro; IPR002939; DnaJ_C.
CC InterPro; IPR001305; DnaJ_CXXCXGKG.
CC InterPro; IPR001623; DnaJ_N.
CC InterPro; IPR003095; Hsp_DnaJ.
CC Pfam; PF00226; DnaJ; 1.
CC Pfam; PF01556; DnaJ_C; 1.
CC Pfam; PF00684; DnaJ_CXXCXGKG; 1.
CC PRINTS; PR00625; DnaJPROTEIN.
CC SMART; SM00271; DnaJ; 1.
CC PROSITE; PS00636; DnaJ_1; FALSE_NEG.
CC PROSITE; PS00076; DnaJ_2; 1.
CC PROSITE; PS00637; DnaJ_CXXCXGKG; FALSE_NEG.
KW Chapterone; DNA replication; Heat shock; Repeat; Zinc; Metal-binding;
KW Complete proteome.
FT DOMAIN 4 68 J-DOMAIN.
FT DOMAIN 79 131 GLY-RICH.
FT REPEAT 165 172 CXXCXGKG MOTIF.
FT REPEAT 182 189 CXXCXGKG MOTIF.
FT REPEAT 208 215 CXXCXGKG MOTIF.
FT REPEAT 222 229 CXXCXGKG MOTIF.
FT METAL 165 168 ZINC 1 (BY SIMILARITY).
FT METAL 168 168 ZINC 1 (BY SIMILARITY).
FT METAL 182 182 ZINC 2 (BY SIMILARITY).
FT METAL 185 185 ZINC 2 (BY SIMILARITY).
FT METAL 208 208 ZINC 2 (BY SIMILARITY).
FT METAL 211 211 ZINC 2 (BY SIMILARITY).
FT METAL 222 222 ZINC 1 (BY SIMILARITY).
FT METAL 225 225 ZINC 1 (BY SIMILARITY).
SQ SEQUENCE 391 AA; 41713 MW; 3FFPD7E01B5EA25F CRC64;

Query Match 1.9%; Score 7; DB 1; Length 391;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
DB 103 GGGAGGG 109
|||||

RESULT 54
PCL_ECTHA STANDARD; PRT; 391 AA.
AC P42516;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 4-coumarate-CoA ligase (EC 6.2.1.12) (4CL) (4-coumaroyl-CoA
DE synthase) (Fragment).
GN PCL.
OS Ectothiorhodospira halophila.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales;
OC Ectothiorhodospiraceae; Halorhodospira.
OX NCBI_TaxID=1053;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BN9626;
RX MEDLINE=96283619; PubMed=8670821;
RA Kort R., Hoff W.D., van West M., Kroon A.R., Hoffer S.M.,
RA Vlieg K.H., Crislaard W., van Beunum J.J., Heilingwerf K.J.;
RT "The xanthopsina: a new family of eubacterial blue-light
RT photoreceptors."
RL EMBO J. 15:3209-3218 (1996).
RN [2]
RP SEQUENCE OF 1-112 FROM N.A.
RC STRAIN=BN9626;
RX MEDLINE=95072006; PubMed=7981196;
RA Baca M., Borgstahl G.E., Boissinot M., Burke P.M., Williams D.R.,

```

```

RA Slater K.A., Getsoff E.D.;
RT "Complete chemical structure of photoactive yellow protein: novel
RT thioester-linked 4-hydroxycinnamyl chromophore and photocycle
RT chemistry."
RL Biochemistry 33:14369-14377 (1994).
CC -|- FUNCTION: CONVERTS P-COUMARIC ACID INTO P-COUMARYL COA. THIS IS
CC NECESSARY FOR THE ACTIVATION OF THE PHOTOACTIVE YELLOW PROTEIN
CC (PYP) CHROMOPHORE.
CC -|- CATALYTIC ACTIVITY: ATP + 4-coumarate + CoA = AMP + diphosphate +
CC 4-coumaroyl-CoA.
CC -|- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL; X98887; CAA67392.1; --
CC EMBL; U17017; AAG61736.1; --
CC PIR; C55993; C55993.
CC InterPro; IPR000873; AMP-bind.
CC Pfam; PF00501; AMP-binding; 2.
CC PROSITE; PS00455; AMP BINDING; FALSE_NEG.
KW Ligase; Phenylpropanoid metabolism.
PT NON_TER 391 391
SQ SEQUENCE 391 AA; 42438 MW; 698321PB6EB2A07D CRC64;

Query Match 1.9%; Score 7; DB 1; Length 391;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 TSGSTGT 85
DB 104 TSGSTGT 110
|||||

RESULT 55
HXAA_HUMAN STANDARD; PRT; 393 AA.
AC P31260; O43370; O43605; Q15949;
DT 01-JUL-1993 (Rel. 26, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Homeobox protein Hox-A10 (Hox-1H) (Hox-1.8) (PL).
GN HOXA10 OR HOX1H.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=91288229; PubMed=1676505;
RA Lowney P., Corral J.C., Detmer K., Lebeau M.M., Deaven L.,
RA Lawrence H.J., Largman C.;
RT "A human Hox 1 homeobox gene exhibits myeloid-specific expression of
RT alternative transcripts in human hematopoietic cells."
RL Nucleic Acids Res. 19:3443-3449 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Mi X., Winters J.L., Stevens D.B., Fleischman R.A.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Bradshaw H., Hinds K., Keppler D.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 301-383 FROM N.A.
RX MEDLINE=90046832; PubMed=2573064;
RA Shen W.-F., Largman C., Lowney P., Corral J.C., Detmer K.,

```

HAUSER C.A., SIMONITCH T.A., HACK F.M., LAWRENCE H.J.;
"Lineage-restricted expression of homeobox-containing genes in human
hematopoietic cell lines";
Proc. Natl. Acad. Sci. U.S.A. 86:8536-8540 (1989).
[5]
SEQUENCE OF 319-384 FROM N.A.
RX MEDLINE=90098876; PubMed=2574852;
RA Acampora D., D'Esposito M., Faiella A., Pannese M., Migliaccio E.,
RA Morelli F., Stornaiuolo A., Nigro V., Simeone A., Boncinelli E.;
"The human HOX gene family";
Nucleic Acids Res. 17:10385-10402 (1989).
[6]
SEQUENCE OF 332-370 FROM N.A.
RX MEDLINE=94145486; PubMed=7906121;
RA Castronovo V., Kusaka M., Charlot A., Gielen J., Sobel M.;
"Homeobox genes: potential candidates for the transcriptional control
of the transformed and invasive phenotype";
Biochem. Pharmacol. 47:137-143 (1994).
CC -!- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH
CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.
CC BINDS TO THE DNA SEQUENCE 5'-AA[AT]TTTATTAC-3'.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=PL1;
CC IsoId=P31260-1; Sequence=Displayed;
CC Name=2; Synonyms=P2;
CC IsoId=P31260-2; Sequence=VSP 002384, VSP 002385;
CC -!- SIMILARITY: BELONGS TO THE ABD-B HOMEBOX FAMILY.
CC -!- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A NUMBER OF
CC FRAMESHIFTS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X58430; CAB86198.1; ALT FRAME.
DR EMBL; AF040714; AAB96917.1; -.
DR EMBL; AC004080; -; NOT ANNOTATED_CDS.
DR EMBL; M30599; AAA36006.1; -.
DR EMBL; S69027; AAD14030.1; -.
DR EMBL; S69029; AAD14031.1; -.
DR PIR; I65265; I65265.
DR PIR; S26402; S26402.
DR PIR; S28814; S28814.
DR HSSP; P02833; 9ANT.
DR TRANSFAC; T01713; -.
DR Genew; HGNC:5100; HOXA10.
DR MIM; 142957; -.
DR GO; GO:0007275; P:development; TAS.
DR GO; GO:0007048; P:oncogenesis; TAS.
DR GO; GO:0007283; P:spermatogenesis; TAS.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation; Alternative splicing.
FT DOMAIN 26 39 POLY-GLY.
FT DOMAIN 78 89 POLY-GLY.
FT DOMAIN 116 120 POLY-PRO.
FT DOMAIN 124 133 POLY-PRO.
FT DOMAIN 208 215 POLY-GLY.
FT DOMAIN 256 259 POLY-PRO.
FT DNA_BIND 319 378 HOMEBOX.

FT VARSPLIC 1 299 Missing (in isoform 2).
FT VARSPLIC 300 302 /FTId=VSP 002384.
FT VARSPLIC 300 302 DSL -> MCQ (in isoform 2).
FT CONFLICT 1 6 MSCSES -> _MCTRVNSQKGLSAPPAKLSHNNVMLGE
FT CONFLICT 69 69 G -> R (IN REF. 1).
FT CONFLICT 136 137 QA -> HR (IN REF. 1).
FT CONFLICT 332 332 L -> S (IN REF. 4).
FT CONFLICT 334 334 L -> F (IN REF. 4).
FT CONFLICT 370 370 R -> P (IN REF. 4).
SQ SEQUENCE 393 AA; 40536 MW; 20F89542582D6F25 CRC64;
Query Match 1.9%; Score 7; DB 1; Length 393;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 235 GCGAGGG 241
DB 29 GCGAGGG 35
|||||
RESULT 56
HEMI_AQUAE STANDARD; PRT; 406 AA.
ID HEMI_AQUAE
AC O67314;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glutamyl-tRNA reductase (EC 1.2.1.-) (Glutr).
GN HEMA OR AQ 1279.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Shead M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus";
RL Nature 392:353-358 (1998).
CC -!- CATALYTIC ACTIVITY: Glutamyl-tRNA(Glu) + NADPH = glutamate-1-
CC semialdehyde + NADP(+) + tRNA(Glu).
CC -!- PATHWAY: Porphyrin biosynthesis by the C5 pathway; first step.
CC -!- SIMILARITY: Belongs to the glutamyl-tRNA reductase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000732; AAC07274.1; -.
DR PIR; G70410; G70410.
DR HSSP; Q42843; IB29.
DR HAMAP; MF 00087; -; 1.
DR InterPro; IPR000343; Glutr.
DR Pfam; PF05201; Glutr_N1.
DR Pfam; PF05200; Glutr_NAD_bind; 1.
DR TIGRFAMs; TIGR01035; hema; 1.
DR PROSITE; PS00747; GLUTR; 1.
KW Porphyrin biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT ACT_SITE 51 51 NUCLEOPHILE (BY SIMILARITY).
FT ACT_SITE 97 97 BASE (BY SIMILARITY).
SQ SEQUENCE 406 AA; 46648 MW; 77F3ECC488FC4394 CRC64;
Query Match 1.9%; Score 7; DB 1; Length 406;
Best Local Similarity 100.0%; Pred. No. 52;

```

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 354 ELAANYL 360
   |||||
Db 193 ELAANYL 199

RESULT 57
YB44_MYCPN
ID YB44_MYCPN STANDARD; PRT; 413 AA.
AC P75142;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Very hypothetical adhesin Pl-like protein MPN144 (E07_orf413):
GN MPN144 OR MP010.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelsreich R., Hilbert H., Plagens H., Pirkle E., Li B.-C.,
Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
pneumoniae."
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- SIMILARITY: BELONGS TO THE ADHESIN P1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000002; AAB95658.1; -.
DR PIR; S73336; S73336.
DR InterPro; IPR004940; Adhesin_P1.
DR Pfam; PF03257; Adhesin_P1; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 413 AA; 43249 MW; 594375273CD5B6E5 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 413;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 235 GGGAGGG 241
   |||||
Db 387 GGGAGGG 393

RESULT 58
SO DROME
ID SO DROME STANDARD; PRT; 416 AA.
AC DQ7350; Q9V4L0;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Sine oculis protein.
GN SO OR C011121.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Head;
RX MEDLINE=95203683; PubMed=7896096;
RA Serikaku M.A., O'Tousa J.E.;

```

```

RT "Sine oculis is a homeobox gene required for Drosophila visual system
development."
RL Genetics 138:1137-1150(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo, and Eye imaginal disk;
RX MEDLINE=94242440; PubMed=7910468;
RA Chevette B.N.R., Green P.J., Martin K., Garren H., Hartenstein V.,
Zipursky S.L.;
RT "The Drosophila sine oculis locus encodes a homeodomain-containing
protein required for the development of the entire visual system."
RL Neuron 12:977-996(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=Berkelley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Anatolides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
Abril J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Miklos G.L.G.,
Bailey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Fabros B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
Foster C., Gabrielian A.E., Garg N.S., Galbart W.M., Glaeser K.,
Glocke A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Hoskin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mout S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: REQUIRED FOR VISUAL SYSTEM DEVELOPMENT. MAY
TRANSCRIPTIONALLY REGULATE GENES NECESSARY FOR OPTIC LOBE
CC -1- INVAGINATION AND BOLWIG'S NERVE FORMATION.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: IN DEVELOPING EMBRYOS, EXPRESSED IN THE EYE
DISK EPITHELIUM, BOLWIG'S ORGAN AND THE OPTIC LOBE PRIMORDIUM AT
AREAS OF INVAGINATION. IN ADULTS, PRESENT IN PHOTORECEPTOR CELLS
IN THE APICAL REGIONS OF THE RETINA, AND IN OPTIC LOBES
CC -1- DEVELOPMENTAL STAGE: IN THE EYE IMAGINAL DISK, FIRST EXPRESSED AT
THE ONSET OF THE THIRD INSTAR AND CONTINUES THROUGHOUT THIS STAGE.
EXPRESSION IN THE OPTIC LOBE PRIMORDIUM BEGINS AT STAGE 5 AND
DISAPPEARS WHEN INVAGINATION IS COMPLETED (STAGE 12). FURTHER
EXPRESSION IS NOTED IN OPTIC LOBE GLANGIA IN LATE THIRD INSTAR.
CC -1- MISCELLANEOUS: SO MUTANTS SPECIFICALLY DISPLAY VISUAL DEFECTS.
THESE RANGE FROM REDUCED OCELLI AND OMATIDIA NUMBER IN WEAK LOSS-
OF-FUNCTION PHENOTYPES TO COMPLETE ABSENCE OF COMPOUND EYES AND
BOLWIG'S ORGAN IN MORE SEVERE LETHAL PHENOTYPES.
CC -1- SIMILARITY: BELONGS TO THE SIX/SINE OCULIS HOMEOBOX FAMILY.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; S77459; AAB34685.1; --
CC EMBL; L31626; AAA21800.1; --
CC EMBL; AE003841; AAF59260.1; --
CC PIR; S50342; S50342.
CC HSSP; P41778; 1DU6.
CC TRNGFAC; T03273; --
CC FlyBase: FBgn0003460; so.
CC GO; GO:0007623; P:circadian rhythm; IGI.
CC GO; GO:0008347; P:glia cell migration; IMP.
CC InterPro; IPR001356; Homeobox.
CC InterPro; IPR000047; HTH lambdarepressor.
CC InterPro; IPR007105; SIX.
CC InterPro; IPR007106; SIX_SINE_homeo.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00031; HTHREPRESSOR.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX 1; 1.
CC PROSITE; PS00071; HOMEBOX_2; 1.
CC Developmental protein; Homeobox; DNA-binding; Nuclear protein.
KW DOMAIN 42 47 POLY-ASN.
FT DOMAIN 52 55 POLY-ASN.
FT DOMAIN 80 85 POLY-GLY.
FT DOMAIN 307 316 POLY-GLN (OPA REPEAT).
FT DOMAIN 366 372 POLY-GLY.
FT DOMAIN 375 379 POLY-ALA.
FT DNA_BIND 218 277 HOMEBOX.
SQ SEQUENCE 416 AA; 45025 MW; 669E4BD95B857996 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 416;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241

Db 65 GGGAGGG 71

RESULT 59

PCL_RHOCA
ID_PCL_RHOCA STANDARD; PRT; 417 AA.
AC O69140;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 4-coumarate--CoA ligase (EC 6.2.1.12) (4CL) (4-coumaroyl-CoA
DE synthase).
GN PCL.
OS Rhodospirillum rubrum (Rhodospirillum rubrum).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Rhodospirillaceae; Rhodospirillum.
OX NCBI_TaxID=1061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SB1003 / St. Louis;
RA Jiang Z., Bauer E.C.;
RT "Genetic characterization of photoactive yellow protein from
RT Rhodospirillum rubrum";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CONVERTS P-COUMARIC ACID INTO P-COUMARYL COA. THIS IS
CC NECESSARY FOR THE ACTIVATION OF THE PHOTOACTIVE YELLOW PROTEIN
CC (PYP) CHROMOPHORE (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + 4-coumarate + CoA = AMP + diphosphate +
CC 4-coumaroyl-CoA.
CC -!- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME

FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AF064095; AAC17429.1; --
CC InterPro; IPR000873; AMP-bind.
CC Pfam; PF00501; AMP-binding; 1.
CC PROSITE; PS00455; AMP BINDING; 1.
KW Ligase; phenylpropanoid metabolism.
SQ SEQUENCE 417 AA; 44047 MW; B3F7BE33D705DBD1 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 417;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 TSGSTGT 85

Db 133 TSGSTGT 139

RESULT 60

CC88_RAT
ID_CC88_RAT STANDARD; PRT; 421 AA.
AC Q8VRW5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Voltage-dependent calcium channel gamma-8 subunit (Neuronal voltage-
DE gated calcium channel gamma-8 subunit).
GN CACNG8.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=21601102; PubMed=11738816;
RA Chu P.-J., Robertson H.M., Best P.M.;
RT "Calcium channel gamma subunits provide insights into the evolution of
RT this gene family";
RL Gene 280:37-48(2001).
CC -!- FUNCTION: THOUGHT TO STABILIZE THE CALCIUM CHANNEL IN AN
CC INACTIVATED (CLOSED) STATE (BY SIMILARITY).
CC -!- SUBUNIT: THE L-TYPE CALCIUM CHANNEL IS COMPOSED OF FIVE SUBUNITS:
CC ALPHA-1, ALPHA-2/DELTA, BETA AND GAMMA.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: BELONGS TO THE PMP-22 / EMP / MP20 FAMILY. CACNG
CC SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AF361346; AAL50041.1; --
CC InterPro; IPR004031; PMP22_Claudin.
CC Pfam; PF00822; PMP22_Claudin; 1.
KW Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
KW Calcium channel.
FT TRANSMEM 19 39 POTENTIAL.
FT TRANSMEM 127 147 POTENTIAL.
FT TRANSMEM 157 177 POTENTIAL.
FT TRANSMEM 207 227 POTENTIAL.

```
FT TRANSMEM 318 338 POTENTIAL.
FT DOMAIN 245 360 GLY-RICH.
SQ SEQUENCE 421 AA; 43268 MW; 0FDB1461BDC780A9 CRC64;

Query Match
Best Local Similarity 1.9%; Score 7; DB 1; Length 421;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGG 240
DB 244 AGGAGG 250
|||||

RESULT 61
CC8_MOUSE STANDARD; PRT; 423 AA.
ID CCG8_MOUSE
AC Q8VHW2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Voltage-dependent calcium channel gamma-8 subunit (Neuronal voltage-
DE sated calcium channel gamma-8 subunit).
GN CACNG8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=21601102; PubMed=11738816;
RA Chu P.-J., Robertson H.M., Best P.M.;
RT "Calcium channel gamma subunits provide insights into the evolution of
RL this gene family.";
RL Gene 280:37-48(2001).
CC -!- FUNCTION: THOUGHT TO STABILIZE THE CALCIUM CHANNEL IN AN
CC INACTIVATED (CLOSED) STATE (BY SIMILARITY).
CC -!- SUBUNIT: THE L-TYPE CALCIUM CHANNEL IS COMPOSED OF FIVE SUBUNITS:
CC ALPHA-1, ALPHA-2/DELTA, BETA AND GAMMA.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: BELONGS TO THE PMP-22 / EMP / MP20 FAMILY. CACNG
CC SUBFAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AF361350; AAL50045.1; -
DR MGD; MGI:1912376; CACNG8
DR InterPro; IPR004031; PMP22_Claudin.
DR Pfam; PF00822; PMP22_Claudin; 1.
KW Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
Calcium channel.
FT TRANSMEM 19 39 POTENTIAL.
FT TRANSMEM 127 147 POTENTIAL.
FT TRANSMEM 157 177 POTENTIAL.
FT TRANSMEM 207 227 POTENTIAL.
FT TRANSMEM 318 338 POTENTIAL.
FT TRANSMEM 245 360 GLY-RICH.
FT DOMAIN 245 360
SQ SEQUENCE 423 AA; 43453 MW; D3AD13E18F9EB02D CRC64;

Query Match
Best Local Similarity 1.9%; Score 7; DB 1; Length 423;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGG 240
DB 244 AGGAGG 250
|||||
```

```
RESULT 62
MKR2_SERQU STANDARD; PRT; 423 AA.
ID MKR2_SERQU
AC Q9DD48;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Makorin 2 (Zinc finger protein YGHLC3HC4).
GN MKR2 OR YGHLC3HC4 OR YGHL2.
OS Seriola quinqueradiata (Five-ray yellowtail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Percoidae;
OC Carangidae; Seriola.
OX NCBI_TaxID=8161;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Takagi R., Saeki T., Abe S.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Gill, and Testis;
RX MEDLINS=21481893; PubMed=11597136;
RA Gray T.A., Azama K., Whitmore K., Min A., Abe S., Nicholls R.D.;
RT "Phylogenetic conservation of the makorin-2 gene, encoding a multiple
RL zine-finger protein, antisense to the raf1 proto-oncogene.";
RL Genomics 77:119-126(2001).
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -!- SIMILARITY: Contains 4 C3H1-type zinc fingers.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AB047517; BAB18861.1; -
DR EMBL; AB049436; BAB18815.1; -
DR EMBL; AB049439; BAB39863.1; -
DR EMBL; AB049437; BAB39861.1; -
DR EMBL; AB049438; BAB39862.1; -
DR HSSP; P28950; 1CHC.
DR InterPro; IPR000571; Znf_CCCH.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR Pfam; PF00642; zf-CCCH; 4.
DR SMART; SM00184; RING; 1.
DR SMART; SM00356; Znf_C3H1; 4.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS0089; ZF_RING_2; 1.
KW Zinc-finger; Repeat.
FT ZN_FING 8 26 C3H1-TYPE 1.
FT ZN_FING 37 55 C3H1-TYPE 2.
FT ZN_FING 168 186 C3H1-TYPE 3.
FT DOMAIN 190 219 MAKORIN-TYPE CYS-HIS.
FT ZN_FING 235 289 RING-TYPE.
FT ZN_FING 324 344 C3H1-TYPE 4.
SQ SEQUENCE 423 AA; 46376 MW; APABF7AB14E86997 CRC64;

Query Match
Best Local Similarity 1.9%; Score 7; DB 1; Length 423;
Matches 7; Conservative 100.0%; Pred. No. 54;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGG 241
DB 74 GGGAGG 80
|||||
```

RESULT 63

CCG8 HUMAN
 ID CCG8 HUMAN STANDARD; PRT; 425 AA.
 AC Q8WX55; Q8BXTO; Q8YX23;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Voltage-dependent calcium channel gamma-8 subunit (Neuronal voltage-gated calcium channel gamma-8 subunit).
 DE CACNG8 OR CACNG6.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 OX [1]
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21601102; PubMed=11738816;
 RA Chu P.-J., Robertson H.M., Best P.M.;
 RT "Calcium channel gamma subunits provide insights into the evolution of this gene family.";
 RT Chromosome 19q13.4: evolution and expression profile of the gamma subunit gene family.";
 RL Gene 280:37-48(2001).
 RN [2]
 RP SEQUENCE OF 12-426 FROM N.A.
 RX MEDLINE=21100909; PubMed=11170751;
 RA Burgess D.B., Geisler L.A., Foreman P.J., Noebels J.L.;
 RT "A cluster of three novel Ca(2+) channel gamma subunit genes on chromosome 19q13.4: evolution and expression profile of the gamma subunit gene family.";
 RL Genomics 71:339-350(2001).
 RN [3]
 RP SEQUENCE OF 1-204 FROM N.A.
 RC TISSUE=Cerebellum;
 RA Black J.L. III, Kryzer T.J., Lennon V.A.;
 RT "Proposed Homo sapiens voltage-gated calcium channel gamma-6 subunit.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THOUGHT TO STABILIZE THE CALCIUM CHANNEL IN AN INACTIVATED (CLOSED) STATE (BY SIMILARITY).
 CC -1- SUBUNIT: THE L-TYPE CALCIUM CHANNEL IS COMPOSED OF FIVE SUBUNITS: ALPHA-1, ALPHA-2/DELTA, BETA AND GAMMA.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE PMP-22 / EMP / MP20 FAMILY. CACNG SUBFAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL; AF361354; AAL50049.1; --
 DR EMBL; AF288388; AAK20031.1; --
 DR EMBL; AF234892; AAK15019.1; ALT_INIT.
 DR Genew; HGNC:13628; CACNG8.
 DR MIM; 606900; --
 DR GO; GO:0005891; C:voltage-gated calcium channel complex; NAS.
 DR GO; GO:0005245; E:voltage-gated calcium channel activity; NAS.
 DR GO; GO:0006816; P:calcium ion transport; NAS.
 DR Pfam; PF00822; PMP22 Claudin; 1.
 DR Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
 KW Calcium channel.
 FT TRANSMEM 19 39
 FT TRANSMEM 129 149
 FT TRANSMEM 158 178
 FT TRANSMEM 208 228
 FT TRANSMEM 318 338
 FT TRANSMEM 247 380
 FT DOMAIN GLY-RICH.
 FT ME -> OV (IN REF. 3).
 FT CONFLICT 1 2
 FT CONFLICT 16 16
 FT CONFLICT 67 70
 FT CONFLICT 362 362
 FT CONFLICT 386 386
 FT CONFLICT 386 386

FT CONFLICT 394 395 RH -> AP (IN REF. 2).
 FT CONFLICT 409 409 G -> E (IN REF. 2).
 SQ SEQUENCE 425 AA; 43481 MW; E11105BDAF619D3D CRC64;
 Query Match 1.9%; Score 7; DB 1; Length 425;
 Best Local Similarity 100.0%; Pred. No. 54;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 234 AGGGAGG 240
 DB 245 AGGGAGG 251
 RESULT 64
 SR54_THEQ STANDARD; PRT; 429 AA.
 ID SR54_THEQ
 AC O07347;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Signal recognition particle protein (Fifty-four homolog).
 DE FFH.
 GN Thermus aquaticus.
 OS Bacteria; Deinococcus-Thermus; Deinococci; Thermaceae;
 OC Thermus.
 OX NCBI_TaxID=271;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Freymann D.M., Keenan R.J., Stroud R.M., Walter P.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS) OF 1-294.
 RX MEDLINE=97156020; PubMed=9002524;
 RA Freymann D.M., Keenan R.J., Stroud R.M., Walter P.;
 RT "Structure of the conserved GTPase domain of the signal recognition particle.";
 RT Nature 385:361-364 (1997).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (3.2 ANGSTROMS).
 RX MEDLINE=98359118; PubMed=9695947;
 RA Keenan R.J., Freymann D.M., Walter P., Stroud R.M.;
 RT "Crystal structure of the signal sequence binding subunit of the signal recognition particle.";
 RT Cell 94:181-191(1998).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.03 ANGSTROMS) OF 1-294.
 RX MEDLINE=99356018; PubMed=10426959;
 RA Freymann D.M., Keenan R.J., Stroud R.M., Walter P.;
 RT "Functional changes in the structure of the SRP GTPase on binding GDP and Mg2+GDP.";
 RT Nat. Struct. Biol. 6:793-801(1999).
 CC -1- FUNCTION: NECESSARY FOR EFFICIENT EXPORT OF EXTRA-CYTOPLASMIC PROTEINS. BINDS TO THE SIGNAL SEQUENCE WHEN IT EMERGES FROM THE RIBOSOMES.
 CC -1- SUBUNIT: SIGNAL RECOGNITION PARTICLE CONSISTS OF A SMALL CYTOPLASMIC RNA (SC-RNA) MOLECULE AND PROTEIN FFH.
 CC -1- DOMAIN: THE PROTEIN HAS A TWO DOMAIN STRUCTURE: THE G-DOMAIN BINDS GTP; THE W-DOMAIN BINDS THE RNA AND ALSO BINDS THE SIGNAL SEQUENCE.
 CC -1- SIMILARITY: BELONGS TO THE SRP FAMILY OF GTP-BINDING PROTEINS.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL; U82109; AAB58502.1; --
 DR PDB; 1FFH; 31-DEC-97.
 DR PDB; 2FFH; 16-JUL-99.
 DR PDB; 1NG1; 13-MAR-00.

DR PDB; 2NG1; 13-MAR-00.
DR PDB; 3NG1; 13-MAR-00.
DR PDB; 1JPU; 02-FEB-02.
DR PDB; 1JPN; 02-FEB-02.
DR PDB; 1LS1; 16-NOV-02.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR000897; SRP54.
DR InterPro; IPR004125; SRP54_SBP.
DR InterPro; IPR004780; SRP_sub.
DR Pfam; PF00448; SRP54; 1.
DR Pfam; PF02881; SRP54_N; 1.
DR Pfam; PF02978; SRP_SBP; 1.
DR ProDom; PD000819; SRP54; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR00959; ffh; 1.
DR PROSITE; PS00300; SRP54; 1.
KW Signal recognition particle; GTP-binding; RNA-binding; 3D-structure.
FT INIT MET 0 0
FT DOMAIN 1 291 G-DOMAIN.
FT DOMAIN 292 429 M-DOMAIN.
FT NP_BIND 104 111 GTP (BY SIMILARITY).
FT NP_BIND 186 190 GTP (BY SIMILARITY).
FT NP_BIND 244 247 GTP (BY SIMILARITY).
FT HELIX 2 13
FT TURN 14 17
FT HELIX 23 39
FT TURN 40 41
FT HELIX 44 59
FT TURN 60 62
FT HELIX 63 65
FT HELIX 69 84
FT TURN 85 86
FT STRAND 98 103
FT TURN 106 107
FT HELIX 110 122
FT TURN 123 125
FT STRAND 128 132
FT HELIX 138 151
FT TURN 152 152
FT STRAND 155 157
FT TURN 160 161
FT HELIX 164 177
FT TURN 178 179
FT STRAND 182 186
FT HELIX 195 208
FT STRAND 212 218
FT HELIX 219 221
FT TURN 222 223
FT HELIX 224 235
FT TURN 236 236
FT STRAND 240 244
FT HELIX 246 248
FT HELIX 253 262
FT STRAND 266 270
FT HELIX 275 277
FT STRAND 278 280
FT HELIX 283 291
FT TURN 292 292
FT HELIX 299 305
FT TURN 321 333
FT TURN 334 334
FT TURN 338 339
FT HELIX 340 342
FT TURN 343 343
FT TURN 351 352
FT HELIX 355 366
FT TURN 367 367
FT HELIX 370 374
FT HELIX 376 378
FT HELIX 381 391
FT TURN 392 392
FT HELIX 395 414
FT TURN 415 416

SQ SEQUENCE 429 AA; 47225 MW; 26D387A5A303EFBD CRC64;
Query Match 1.9%; Score 7; DB 1; Length 429;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 326 QEAIGRL 332
DB 9 QEAIGRL 15
RESULT 65
ELT2_CABEL
ID ELT2_CABEL STANDARD; PRT; 433 AA.
AC Q10655; Q18371;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Transcription factor elt-2.
GN ELT-2 OR C33D3.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=95301560; PubMed=7782329;
RA Hawkins M.G., McGhee J.D.;
RT "elt-2, a second GATA factor from the nematode Caenorhabditis elegans.";
RL J. Biol. Chem. 270:14666-14671(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Sulston J.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP REVISIONS.
RA Sulston J.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: TRANSCRIPTIONAL ACTIVATOR OF GENES CONTAINING THE GATA REGION.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DEVELOPMENTAL STAGE: PRESENT AT ALL STAGES BUT HIGHEST IN EMBRYOS.
CC -!- SIMILARITY: Contains 1 GATA-type zinc finger.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
DR EMBL; U25175; AAC316130.1; -
DR EMBL; Z49867; CA930029.2; -
DR FIR; A56953; A56953.
DR HSSP; P17678; 1GAT.
DR WormPep; C33D3.1; CE31430.
DR InterPro; IPR000679; Znf_GATA.
DR Pfam; PF00320; GATA; 2.
DR PRINTS; PR00619; GATAZNFINGER.
DR SMART; SM00401; Znf_GATA; 1.
DR PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
DR PROSITE; PS0114; GATA_ZN_FINGER_2; 1.
KW Transcription regulation; Activator; DNA-binding; Zinc-finger;
KW Metal-binding; Nuclear protein.
FT ZN_FING 237 261 GATA-TYPE.
SQ SEQUENCE 433 AA; 47116 MW; 59C2DEB0753E5499 CRC64;
Query Match 1.9%; Score 7; DB 1; Length 433;

Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 183 LYSIGIPV 189
105 LYSIGIPV 111

Db

RESULT 66
K1CM MOUSE
ID K1CM MOUSE STANDARD; PRT; 437 AA.
AC P08730;
DT 01-AUG-1998 (Rel. 08, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Keratin, type I cytoskeletal 13 (Cytokeratin 13) (K13) (CK 13) (47 kDa
cytokeratin).
GN KRT13 OR KRT13-13.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=95213021; PubMed=7535287;
RA Fillon M., Sarafian V., Lussier M., Belanger C., Lapointe L.,
RA Royal A.;
RT "Arrangement of a cluster of three mouse type I keratin genes
expressed sequentially during esophageal-type epithelial cell
differentiation";
RL Genomics 24:303-310 (1994).
RN [2]
RP SEQUENCE OF 126-437 FROM N.A.
RX MEDLINE=86120369; PubMed=2418416;
RA Knapp B., Rentrop M., Schweizer J., Winter H.;
RT "Nonepidermal members of the keratin multigene family: cDNA sequences
and in situ localization of the mRNAs";
RL Nucleic Acids Res. 14:751-763 (1986).
RN [3]
RP SEQUENCE OF 1-157 FROM N.A.
RX MEDLINE=90323435; PubMed=1695590;
RA Winter H., Rentrop M., Nischt R., Schweizer J.;
RT "Tissue-specific expression of murine keratin K13 in internal
stratified squamous epithelia and its aberrant expression during
two-stage mouse skin carcinogenesis is associated with the
methylation state of a distinct CpG site in the remote 5'-flanking
region of the gene";
RL Differentiation 43:105-114 (1990).
CC -1- SUBUNIT: HETEROTETRAMER OF TWO TYPE I AND TWO TYPE II KERATINS.
CC -1- MISCELLANEOUS: There are two types of cytoskeletal and
microfibrillar keratin, I (acidic) and II (neutral to basic) (40-
55 and 56-70 kDa, respectively).
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U13921; AAC52150.1; -;
CC EMBL; X03492; CAA27208.1; -;
CC EMBL; X53320; CAA37407.1; -;
CC PIR; A55682; A55682.
CC MGI; MGI:101925; Krt1-13.
CC InterPro; IPR001664; IF.
CC InterPro; IPR002957; Keratin_I.
CC Pfam; PF00038; filament_1.
CC PRINTS; PR01248; TYPE1KERATIN.
CC PROSITE; PS00226; IF; 1.

Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 GGGAGGG 241
54 GGGAGGG 60

Db

RESULT 67
FXGA CHICK
ID FXGA CHICK STANDARD; PRT; 440 AA.
AC Q98937;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Forkhead box protein G1A (Forkhead-related protein FKHL2)
DE (Transcription factor BF-2) (Brain factor 2) (BF2) (CBF-2) (T-14-6).
GN FOXG1A OR FKHL2 OR HFHBF2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=White leghorn; TISSUE=Retina;
RX MEDLINE=96338226; PubMed=8757134;
RA Yuasa J., Hirano S., Yamagata M., Noda M.;
RT "Visual projection map specified by topographic expression of
transcription factors in the retina";
RL Nature 382:632-635 (1996).
CC -1- FUNCTION: MAY DETERMINE THE NASOTEMPORAL AXIS OF THE RETINA, AND
CONSEQUENTLY SPECIFY THE TOPOGRAPHICAL PROJECTION OF THE RETINAL
GANGLION-CELL AXONS TO THE TECTUM BY CONTROLLING EXPRESSION OF
THEIR TARGET GENES.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- TISSUE SPECIFICITY: Retina.
CC -1- DEVELOPMENTAL STAGE: CAN BE DETECTED IN REGIONS INCLUDING
PRIMORDIAL RETINA AND NEUROEPITHELIUM BY EMBRYONIC DAY 2 (E2). AT
E3, EXPRESSED IN THE TEMPORAL RETINA AND ASSOCIATED PIGMENT
EPITHELIUM AS WELL AS IN PART OF THE DIENCEPHALON, AND AT E7 IS
EXPRESSED IN RETINAL GANGLION CELLS. LEVELS BEGIN TO DECLINE FROM
E4 AND ALMOST DISAPPEAR BY E10.
CC -1- SIMILARITY: Contains 1 fork-head domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U47276; AAB08467.1; -;
CC PIR; S71795; S71795.
CC HSSP; Q63245; 2HFH.
CC InterPro; IPR001766; TF Fork head.
CC Pfam; PF00250; Fork head; 1.
CC PRINTS; PR00053; FORKHEAD.
CC ProDom; PD000425; TF Fork_head; 1.
CC SMART; SM00339; FH; 1.

DR PROSITE; PS00657; FORK HEAD 1; 1.
 DR PROSITE; PS00658; FORK HEAD 2; 1.
 DR PROSITE; PS00339; FORK HEAD 3; 1.
 KW Transcription regulation; DNA-binding; Nuclear protein;
 FT Developmental protein.
 FT DOMAIN 39 43 POLY-ARG.
 FT DOMAIN 49 58 POLY-GLU.
 FT DOMAIN 77 88 POLY-GLY.
 FT DOMAIN 109 112 POLY-ALA.
 FT DOMAIN 113 137 POLY-GLY.
 FT DNA BIND 143 237 FORK-HEAD.
 FT DOMAIN 321 330 POLY-PRO.
 SQ SEQUENCE 440 AA; 44669 MW; 82266C3E1E103A48 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 440;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
 DB 86 GGGAGGG 92
 |||||

RESULT 68

SOX4_MOUSE
 ID SOX4_MOUSE STANDARD; PRT; 440 AA.
 AC Q06831;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Transcription factor SOX-4.
 OS SOX4 OR SOX-4.
 GN Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94008992; PubMed=8404853;
 RA van de Wetering M., Oosterwegel M., van Norren K., Clevers H.C.;
 RT "Sox-4, an Sry-like HMG box protein, is a transcriptional activator
 in lymphocytes.";
 RL EMBO J. 12:3847-3854 (1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93261838; PubMed=8493110;
 RA Schilham M.W., van Eijk M., van de Wetering M., Clevers H.C.;
 RT "The murine Sox-4 protein is encoded on a single exon.";
 RL Nucleic Acids Res. 21:2009-2009 (1993).
 RN [3]

RP SEQUENCE OF 69-122 FROM N.A.
 RC STRAIN=Swiss Webster; TISSUE=Embryonic tooth;
 RX MEDLINE=97079683; PubMed=8921394;
 RA Stock D.W., Buchanan A.V., Zhao Z., Weiss K.M.;
 RT "Numerous members of the Sox family of HMG box-containing genes are
 expressed in developing mouse teeth.";
 RL Genomics 37:234-237 (1996).
 CC -1- FUNCTION: TRANSCRIPTIONAL ACTIVATOR THAT BINDS WITH HIGH AFFINITY
 TO THE T-CELL ENHANCER MOTIF 5'-AACAAAG-3' MOTIF.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN LYMPHOCYTES AND IN MOLAR AND
 INCISOR TOOTH GERMS.
 CC -1- SIMILARITY: Contains 1 HMG box domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
 or send an email to license@isb-sib.ch).

DR EMBL; X70298; CAA49779.1; -.

DR EMBL; U70440; AAC52858.1; -.
 DR PIR; S37303; S37303.
 DR HSSP; P48436; 1SX9.
 DR TRANSFAC; T01838; -.
 DR MGD; MGI:98366; Sox4.
 DR InterPro; IPR000910; HMG 12_box.
 DR Pfam; PF00505; HMG_box; 1.
 DR SMART; SM00398; HMG; 1.
 DR PROSITE; PS01118; HMG BOX 2; 1.
 KW DNA-binding; Nuclear protein; Transcription regulation; Activator.
 FT DNA BIND 59 127 HMG BOX.
 FT DOMAIN 347 363 POLY-SER.
 SQ SEQUENCE 440 AA; 45187 MW; 60BBE7FBF6FFBC21 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 440;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGGG 240
 DB 168 AGGAGGG 174
 |||||

RESULT 69

SOX11_HUMAN
 ID SOX11_HUMAN STANDARD; PRT; 441 AA.
 AC P35716;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Transcription factor SOX-11.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96115609; PubMed=8666406;
 RA Jay P., Goze C., Marsollier C., Taviaux S., Hardelin J.-P.,
 RA Koopman P., Berta P.;
 RT "The human SOX11 gene: cloning, chromosomal assignment and tissue
 expression.";
 RL Genomics 29:541-545 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20039622; PubMed=10574465;
 RA Azuma T., Ao S., Saito Y., Yano K., Seki N., Wakao H., Masuho Y.,
 RA Muramatsu M.;
 RT "Human SOX11, an upregulated gene during the neural differentiation,
 has a long 3' untranslated region.";
 RL DNA Res. 6:357-360 (1999).
 RN [3]

RP SEQUENCE OF 57-115 FROM N.A.
 RX MEDLINE=93324385; PubMed=8332506;
 RA Goze C., Poulat P., Berta P.;
 RT "Partial cloning of SOX-11 and SOX-12, two new human SOX genes.";
 RL Nucleic Acids Res. 21:2943-2943 (1993).
 CC -1- FUNCTION: PROBABLY IMPORTANT IN THE DEVELOPING NERVOUS SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- TISSUE SPECIFICITY: EXPRESSED MAINLY IN THE NERVOUS SYSTEM.
 CC -1- SIMILARITY: Contains 1 HMG box domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
 or send an email to license@isb-sib.ch).

DR EMBL; U23752; AAB08518.1; -.

```
DR EMBL; AB028641; BAA88122.1; -.
DR EMBL; X73038; CAA51519.1; -.
DR PIR; G01758; G01758.
DR PIR; S34118; S34118.
DR HSP; P48436; I8X9.
DR TRANSFAC; T02894; -.
DR Genew; HGNC:11191; SOX11.
DR MIM; 600898; -.
DR GO; GO:0007399; P:neurogenesis; TAS.
DR InterPro; IPR000910; HMG_12_box.
DR Pfam; PF00505; HMG_box; 1.
DR SMART; SM00398; HMG; 1.
DR PROSITE; PS0118; HMG_BOX 2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein.
FT DNA BIND 49 117 HMG_BOX
FT DOMAIN 144 148 POLY-GLY.
FT DOMAIN 207 214 POLY-GLY.
FT DOMAIN 223 233 POLY-ASP.
FT DOMAIN 338 354 POLY-SER.
SQ SEQUENCE 441 AA; 46679 MW; 50567A654F503C63 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 441;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGG 240
Db 150 AGGGAGG 156
|||||

RESULT 70
COBB METJA
ID _COBB METJA STANDARD; PRT; 443 AA.
AC QS8816;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable cobyrinic acid A,C-diamide synthase.
GN COBB OR MJ1421
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=868087;
RA Buit C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-P., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii";
RL Science 273:1058-1073(1996).
CC -!- FUNCTION: RESPONSIBLE FOR THE AMIDATION OF CARBOXYLIC GROUPS AT
CC POSITION A AND C OF EITHER COBYRINIC ACID OR HYDROGENOBYRINIC ACID.
CC NH(2) GROUPS ARE PROVIDED BY GLUTAMINE, AND ONE MOLECULE OF ATP
CC IS HYDROGENOLYZED FOR EACH AMIDATION (BY SIMILARITY).
CC -!- PATHWAY: Cobalamin biosynthesis.
CC -!- SIMILARITY: Belongs to the cobB/cobQ family. CobB subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67583; AAB99432.1; -.
DR PIR; D64477; D64477.
DR TIGR; MJ1421; -.
DR HAMAP; MF_00027; -; 1.
DR InterPro; IPR004484; CbiA_P.
DR InterPro; IPR002586; CbiA_P.
DR Pfam; PF01656; CbiA; 1.
DR TIGRFAMs; TIGR00379; cobB; 1.
KW Cobalamin biosynthesis; Porphyrin biosynthesis; Complete proteome.
SQ SEQUENCE 443 AA; 50205 MW; EA8D6D0FC9ED212 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 443;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 47 IFNGKVL 53
Db 414 IFNGKVL 420
|||||

RESULT 71
GLNA HALVO
ID _GLNA HALVO STANDARD; PRT; 454 AA.
AC P43386;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Glutamine synthetase (EC 6.3.1.2) (Glutamate--ammonia ligase) (GS).
GN GLNA.
OS Halobacterium volcanii (Haloflex volcanii).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Haloflex.
OX NCBI_TaxID=2246;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DS2;
RX MEDLINE=94365840; PubMed=7916055;
RA Brown J.R., Masuchi Y., Robb F.T., Doolittle W.F.;
RA "Evolutionary relationships of bacterial and archaeal glutamine
RA synthetase genes";
RT J. Mol. Evol. 38:566-576(1994).
CC -!- CATALYTIC ACTIVITY: ATP + L-glutamate + NH(3) = ADP + phosphate +
CC L-glutamine.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U03029; AAC43489.1; -.
DR PIR; T47127; T47127.
DR HSP; P06201; ILGR.
DR InterPro; IPR001691; GLN_synth.
DR InterPro; IPR004809; GLNA.
DR InterPro; IPR001637; GLNA_adenyltn.
DR Pfam; PF00120; gln-synt; 1.
DR Pfam; PF03951; gln-synt N; 1.
DR ProDom; PD001057; gln_synt C; 1.
DR TIGRFAMs; TIGR00653; GLNA_1.
DR PROSITE; PS00180; GLNA_1; 1.
DR PROSITE; PS00181; GLNA_ATP; 1.
DR BINDING 384 384 AMP (UNDER CONDITIONS OF ABUNDANT
DR PROSITE; PS00181; GLNA_ATP; 1.
FT BINDING 384 384 AMP (UNDER CONDITIONS OF ABUNDANT
FT PROSITE; PS00181; GLNA_ATP; 1.
SQ SEQUENCE 454 AA; 50696 MW; F07ECC3133B07722 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 454;
Best Local Similarity 100.0%; Pred. No. 58;
```

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 LESMGFD 338
 Db 186 LESMGFD 192

RESULT 72
 FXD1_HUMAN STANDARD; PRT; 465 AA.
 AC Q16676; Q12949;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Forkhead box protein D1 (Forkhead-related protein FKHL8) (Forkhead-
 DE related transcription factor 4) (FREAC-4).
 GN FOXD1 OR FKHL8 OR FREAC4.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96355487; PubMed=8702877;
 RA Ernersson S., Pierrou S., Hulander M., Cederberg A., Hellqvist M.,
 RA Carlsson P., Enerback S.;
 RT "Characterization of the human forkhead gene FREAC-4. Evidence for
 RT regulation by Wilms' tumor suppressor gene (WT-1) and p53.";
 RL J. Biol. Chem. 271:21094-21099(1996).
 RN [2]
 RP SEQUENCE OF 120-225 FROM N.A.
 RX MEDLINE=95045392; PubMed=7957066;
 RA Pierrou S., Hellqvist M., Samuelsson L., Enerbaeck S., Carlsson P.;
 RT "Cloning and characterization of seven human forkhead proteins:
 RT binding site specificity and DNA bending.";
 RL EMBO J. 13:5002-5012(1994).
 CC -1- FUNCTION: BINDING OF FREAC-3 AND FREAC-4 TO THEIR COGNATE SITES
 CC RESULTS IN BENDING OF THE DNA AT AN ANGLE OF 80-90 DEGREES.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Contains 1 fork-head domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U59832; AAC50661.1; -;
 DR EMBL; U59831; AAC50660.1; -;
 DR EMBL; U13222; AAA92039.1; -;
 DR PIR; G02738; G02738.
 DR PIR; S51627; S51627.
 DR HSP; Q63245; 2HFH.
 DR TRANSFAC; T02472; -;
 DR Genew; HGNC:3802; FOXD1.
 DR MTM; 601091; -;
 DR GO; GO:0003700; F:transcription factor activity; TAS.
 DR InterPro; IPR001766; TF_Fork_head.
 DR Pfam; PF00250; Fork_head; 1.
 DR PRINTS; PR00053; FORKHEAD.
 DR ProDom; PD000425; TF_Fork_head; 1.
 DR SMART; SM00339; FH; 1.
 DR PROSITE; PS00657; FORK_HEAD_1; 1.
 DR PROSITE; PS00658; FORK_HEAD_2; 1.
 DR PROSITE; PS50039; FORK_HEAD_3; 1.
 KW Transcription regulation; DNA-binding; Nuclear protein.
 FT DOMAIN 26 34 POLY-GLU.
 FT DOMAIN 39 43 POLY-GLY.
 FT DOMAIN 52 57 POLY-ARG.
 FT DOMAIN 69 72 POLY-GLU.
 FT DOMAIN 73 76 POLY-ASP.

FT DOMAIN 97 113 POLY-GLY.
 FT DNA_BIND 124 215 FORK-HEAD.
 FT DOMAIN 231 234 POLY-ALA.
 FT DOMAIN 252 256 POLY-ALA.
 FT DOMAIN 259 266 POLY-PRO.
 FT DOMAIN 293 303 POLY-ALA.
 FT DOMAIN 309 315 POLY-PRO.
 FT DOMAIN 375 390 POLY-ALA.
 FT DOMAIN 428 434 POLY-ALA.
 SQ SEQUENCE 465 AA; 46140 MW; D3E7854909CCBFAC CRC64;

Query Match 1.9%; Score 7; DB 1; Length 465;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
 Db 106 GGGAGGG 112

RESULT 73
 FXD3_MOUSE STANDARD; PRT; 465 AA.
 AC Q61660;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Forkhead box protein D3 (HNF3/FH transcription factor genesis)
 DE (Hepatocyte nuclear factor 3 forkhead homolog 2) (HFH-2).
 GN FOXD3 OR HFH2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hromas R.A., Costa R.H., Xu D., Sutton J.L.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Labosky P.A., Kaetner K.H.;
 RT "The winged helix transcription factor Hfh2 is expressed in neural
 RT crest and spinal cord during mouse development.";
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PROBABLE TRANSCRIPTION FACTOR.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Contains 1 fork-head domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U41047; AAA87569.1; -;
 DR EMBL; AF067421; AAC28352.1; -;
 DR HSP; Q63245; 2HFH.
 DR TRANSFAC; T04166; -;
 DR MGD; MGI:1347473; Foxd3.
 DR InterPro; IPR001766; TF_Fork_head.
 DR Pfam; PF00250; Fork_head; 1.
 DR PRINTS; PR00053; FORKHEAD.
 DR ProDom; PD000425; TF_Fork_head; 1.
 DR SMART; SM00339; FH; 1.
 DR PROSITE; PS00657; FORK_HEAD_1; 1.
 DR PROSITE; PS00658; FORK_HEAD_2; 1.
 DR PROSITE; PS50039; FORK_HEAD_3; 1.
 KW DNA-binding; Nuclear protein; Transcription regulation.
 FT DOMAIN 106 115 POLY-GLY.
 FT DNA_BIND 131 225 FORK-HEAD.
 FT DOMAIN 252 257 POLY-ALA.

```
FT DOMAIN 265 270 POLY-ALA.
FT DOMAIN 275 281 POLY-ALA.
FT DOMAIN 380 399 POLY-GLY.
FT DOMAIN 447 457 POLY-ALA.
SQ SEQUENCE 465 AA; 47092 MW; 6F8B5B3D8C7564D CRC64;

Query Match 1.9%; Score 7; DB 1; Length 465;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 GGGAGGG 241
|||||
Db 385 GGGAGGG 391

RESULT 74
SYN THETN STANDARD; PRT; 467 AA.
AC Q8RT3;
DT 28-FEB-2003 (Rel. 41, Last Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cysteine--trna synthetase (EC 6.1.1.16) (Cysteine--trna ligase)
DE (CYERS).
GN CYSS OR TTE2315.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
CC -!- CAVALYTIC ACTIVITY: ATP + L-cysteine + tRNA(Cys) = AMP +
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC Strong, to methionyl-tRNA synthetase.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AE013173; AM25456.1; -.
DR HAMAP; MF_00041; -.
DR InterPro; IPR002308; Cys trna-synt 1a.
DR InterPro; IPR001412; trna-synt 1.
DR Pfam; PF01406; trna-synt 1e; 1.
DR TIGRFAMs; TIGR00435; cy85; 1.
DR PROSITE; PS00178; AA tRNA LIGASE I; FALSE NEG.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 29 39 "HIGH" REGION.
FT SITE 264 268 "KMSKS" REGION.
FT BINDING 267 267 ATP (BY SIMILARITY).
SQ SEQUENCE 467 AA; 54523 MW; B836F5F0CD82BF6B CRC64;

Query Match 1.9%; Score 7; DB 1; Length 467;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 280 IRLIEE 286
|||||
```

```
Db 421 IRLIEE 427

RESULT 75
K1C0_XENLA STANDARD; PRT; 486 AA.
ID K1C0_XENLA
AC P02537;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Keratin 3, type I cytoskeletal 51 kDa (51 kDa cyokeratin).
OS Xenopus laevis (African clawed frog).
OC Sauriata; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89005622; PubMed=2458965;
RA Hoffmann W., Sterrer S., Koenigstorfer A.;
RT "Amino acid sequence microheterogeneities of a type I cyokeratin of
RT Mr 51,000 from Xenopus laevis epidermis.";
RL FEBS Lett. 237:178-182(1988).
RN [2]
RP SEQUENCE OF 339-386 FROM N.A.
RX MEDLINE=84261417; PubMed=6204859;
RA Hoffmann W., Franz J.K.;
RT "Amino acid sequence of the carboxy-terminal part of an acidic type I
RT cyokeratin of molecular weight 51 000 from Xenopus laevis epidermis
RT as predicted from the cDNA sequence.";
RL EMBO J. 3:1301-1306(1984).
CC -!- SUBUNIT: HETEROTETRAMER OF TWO TYPE I AND TWO TYPE II KERATINS.
CC -!- MISCELLANEOUS: There are two types of cytoskeletal and
CC microfibrillar keratin, I (acidic) and II (neutral to basic) (40-
CC 55 and 56-70 kDa, respectively).
CC -!- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; Y00968; CAA68783.1; -.
DR PIR; S01327; KRXL.
DR InterPro; IPR001664; IF.
DR InterPro; IPR002957; Keratin_I.
DR Pfam; PF00038; Filament; 1.
DR PRINTS; PR01248; TYPE1KERATIN.
DR PROSITE; PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Keratin.
FT DOMAIN 1 125 HEAD.
FT DOMAIN 126 438 ROD.
FT DOMAIN 439 486 TAIL.
FT DOMAIN 126 161 COIL 1A.
FT DOMAIN 162 184 LINKER 1.
FT DOMAIN 185 276 COIL 1B.
FT DOMAIN 277 299 LINKER 12.
FT DOMAIN 300 438 COIL 2.
FT SITE 380 380 STUTTER.
FT CONFLICT 352 352 R -> T (IN REF. 2).
SQ SEQUENCE 486 AA; 51889 MW; 9B6D026C597109C5 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 486;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 GGGAGGG 241
|||||
Db 70 GGGAGGG 76
```

Search completed: December 17, 2003, 06:33:52
Job time : 14.4735 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 06:26:06 ; Search time 36.1811 Seconds
(without alignments)
2624.667 Million cell updates/sec

Title: US-09-805-550-4

Perfect score: 388

Sequence: 1 MKLTVTLKGTHTFEIRVQPN.....CDRNEELAANYLLEHAGED 368

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

SPTREMBL 23.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	5.7	365	10 Q94CE9	Q94CE9 arabidopsis
2	21	5.7	367	10 Q9MA10	Q9MA10 arabidopsis
3	21	5.7	371	10 Q8LA46	Q8LA46 arabidopsis
4	15	4.1	113	10 Q9SA20	Q9SA20 arabidopsis
5	15	4.1	378	10 Q94C35	Q94C35 arabidopsis
6	15	4.1	378	10 Q9FF16	Q9FF16 arabidopsis
7	15	4.1	379	10 Q03991	Q03991 daucus caro
8	15	4.1	382	10 Q03990	Q03990 daucus caro
9	14	3.8	246	10 Q9S9L8	Q9S9L8 arabidopsis
10	14	3.8	389	10 Q9STA6	Q9STA6 lycopersico
11	13	3.5	392	10 Q40742	Q40742 oryza sativ
12	11	3.0	343	5 Q8IMB7	Q8IMB7 drosophila
13	11	3.0	414	5 Q9XZ50	Q9XZ50 drosophila
14	11	3.0	414	5 Q9V3W9	Q9V3W9 drosophila
15	11	3.0	419	10 Q9M887	Q9M887 arabidopsis
16	10	2.7	65	10 Q9SCA8	Q9SCA8 lycopersico

9	2.4	225	16	Q9RU37	Q9RU37 deinococcus
9	2.4	753	10	Q9C552	Q9C552 arabidopsis
8	2.2	111	10	Q39682	Q39682 daucus caro
8	2.2	125	11	Q8CA15	Q8CA15 mus musculus
8	2.2	139	10	Q9FQY9	Q9FQY9 nicotiana t
21	8	167	11	Q924J3	Q924J3 tamiag stri
22	8	173	10	Q41191	Q41191 arabidopsis
23	8	174	10	Q9LTP5	Q9LTP5 arabidopsis
24	8	186	3	Q96V99	Q96V99 pneumocysti
25	8	202	10	Q9LGP0	Q9LGP0 oryza sativ
26	8	208	10	Q8LA18	Q8LA18 oryza sativ
27	8	210	10	Q41187	Q41187 arabidopsis
28	8	220	10	Q9367	Q9367 brassica ol
29	8	228	11	Q8C4G2	Q8C4G2 mus musculus
30	8	228	11	Q8BPM6	Q8BPM6 mus musculus
31	8	307	10	Q8S7T7	Q8S7T7 oryza sativ
32	8	340	16	Q9KYD8	Q9KYD8 streptomyce
33	8	360	11	Q9Z2C8	Q9Z2C8 mus musculus
34	8	366	6	Q9GLB8	Q9GLB8 cephalorhyn
35	8	373	3	Q9P639	Q9P639 neurospora
36	8	384	11	Q9ROJ6	Q9ROJ6 apodemus sp
37	8	384	11	Q9ROH6	Q9ROH6 rattus norv
38	8	384	11	Q9ROJ5	Q9ROJ5 apodemus pe
39	8	384	11	Q9ROJ4	Q9ROJ4 apodemus ag
40	8	384	11	Q9ROH7	Q9ROH7 tokudaia os
41	8	387	10	Q8GVV7	Q8GVV7 oryza sativ
42	8	390	10	Q9M3Y2	Q9M3Y2 triticum ae
43	8	395	11	Q64217	Q64217 spalax zemn
44	8	400	10	Q8GTL0	Q8GTL0 oryza sativ
45	8	401	10	Q9C518	Q9C518 arabidopsis
46	8	420	10	Q8H010	Q8H010 oryza sativ
47	8	423	11	Q8X1Q2	Q8X1Q2 rattus norv
48	8	424	11	Q8X1P1	Q8X1P1 tachyorycte
49	8	427	5	Q9GR85	Q9GR85 halocynthia
50	8	438	5	Q9NG13	Q9NG13 branchiost
51	8	458	16	Q8EJY7	Q8EJY7 xanthomonas
52	8	458	16	Q8P8G0	Q8P8G0 xanthomonas
53	8	492	4	Q9UNW9	Q9UNW9 homo sapien
54	8	492	11	Q35392	Q35392 mus musculus
55	8	498	4	Q43267	Q43267 homo sapien
56	8	544	5	Q8I6C7	Q8I6C7 drosophila
57	8	544	5	Q8I6C6	Q8I6C6 drosophila
58	8	544	5	Q8I6C5	Q8I6C5 drosophila
59	8	544	5	Q8I6C4	Q8I6C4 drosophila
60	8	544	5	Q8I6C3	Q8I6C3 drosophila
61	8	544	5	Q8I6B7	Q8I6B7 drosophila
62	8	544	5	Q8I072	Q8I072 drosophila
63	8	548	5	Q22809	Q22809 caenorhabdi
64	8	550	5	Q8IGB4	Q8IGB4 drosophila
65	8	556	5	Q9VTJ7	Q9VTJ7 drosophila
66	8	579	10	Q8LRM4	Q8LRM4 chlamydomon
67	8	579	10	Q8LRM5	Q8LRM5 chlamydomon
68	8	584	16	Q9FCJ3	Q9FCJ3 streptomyce
69	8	712	5	Q43981	Q43981 eimeria ten
70	8	731	10	Q9AMU2	Q9AMU2 oryza sativ
71	8	738	5	O02402	O02402 pinctada fu
72	8	795	5	O62006	O62006 branchiost
73	8	809	10	Q43715	Q43715 pisum sativ
74	8	823	5	Q25343	Q25343 teishmania
75	8	968	5	Q9VKL0	Q9VKL0 drosophila
76	8	1004	16	Q8NP40	Q8NP40 corynebacte
77	8	1010	2	Q45340	Q45340 bordetella
78	8	1014	5	Q9V9Q9	Q9V9Q9 drosophila
79	8	1234	11	Q8VD34	Q8VD34 mus musculus
80	8	1333	5	Q9V9Q7	Q9V9Q7 drosophila
81	8	1385	5	Q8T0L7	Q8T0L7 drosophila
82	8	1385	5	Q9VVC7	Q9VVC7 drosophila
83	8	1457	5	Q8SY47	Q8SY47 drosophila
84	8	1718	5	O62603	O62603 trypanosoma
85	8	2075	5	Q9VXY2	Q9VXY2 drosophila
86	8	2425	17	O28859	O28859 archaeglob
87	8	2904	11	Q9EPN0	Q9EPN0 mus musculus
88	8	2931	11	Q9EPM9	Q9EPM9 mus musculus

90 Q9epn1 mus musculus
 91 P73590 synchocyst
 92 Q9trd7 oryctolagus
 93 Q23471 arabidopsis
 94 Q98857 glycine max
 95 Q9wur8 rattus norv
 96 Q9uki4 homo sapien
 97 Q93ka3 erwinia chr
 98 Q50370 mycoplasma
 99 Q8nkr0 drosophila
 100 Q8h778 arabidopsis

ALIGNMENTS

RESULT 1
 ID Q94CE9 PRELIMINARY; PRT; 365 AA.
 AC Q94CE9;
 DT 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Putative RAD23 protein.
 GN F20817.8.

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,
 RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
 RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koeseema E.,
 RA Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,
 RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "Full Length cDNA of gene F20817.8 (GI:7715605).";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Banno P., Chang E., Dale J.M., Goldsmith A.D.,
 RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
 RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Kim C., Koeseema E., Lam B., Lin J., Meyers M.C., Miranda M.,
 RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
 RA Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.;
 RT "Full Length cDNA of gene F20817.8 (GI:7715605).";
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY034912; AAK59419.1; -;
 DR EMBL; AY063103; AAL34277.1; -;
 DR InterPro; IPR004806; Rad23.
 DR InterPro; IPR006636; ST11.
 DR InterPro; IPR000449; UBA domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 2.
 DR Pfam; PF00240; ubiquitin; 1.
 DR SMART; SM00727; ST11; 1.
 DR SMART; SM00165; UBA; 2.
 DR SMART; SM00213; UBQ; 1.
 DR TIGRFAMs; TIGR00601; rad23; 1.
 DR PROSITE; PS50053; UBIQUITIN_2; 1.
 SQ SEQUENCE 365 AA; 39157 MW; 4FED9EC598467745 CRC64;

Query Match 5.7%; Score 21; DB 10; Length 365;
 Best Local Similarity 100.0%; Pred. No. 2.5e-12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 342 VIEAPLACDRNEELAANYLLE 362

Db 338 VIEAPLACDRNEELAANYLLE 358

RESULT 2

ID Q9MA10 PRELIMINARY; PRT; 367 AA.
 AC Q9MA10;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE F20817.8.

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Khan S., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Kim C.,
 RA Shinn P., Altafi H., Bei Q., Chin C., Chiou J., Choi E., Conn L.,
 RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
 RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
 RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
 RA Ecker J.R.;
 RT "Genomic sequence for Arabidopsis thaliana BAC F20817 from chromosome
 I-1";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
 RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E.,
 RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
 RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
 RA Theologis A., Ecker J.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC010793; AAF68123.1; -;
 DR HSSP; P54725; 1DV0.
 DR InterPro; IPR004806; Rad23.
 DR InterPro; IPR000449; UBA domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 2.
 DR Pfam; PF00240; ubiquitin; 1.
 DR SMART; SM00165; UBA; 2.
 DR SMART; SM00213; UBQ; 1.
 DR TIGRFAMs; TIGR00601; rad23; 1.
 DR PROSITE; PS50053; UBIQUITIN_2; 1.
 SQ SEQUENCE 367 AA; 39790 MW; C6D01134B5C69CE6 CRC64;

Query Match 5.7%; Score 21; DB 10; Length 367;
 Best Local Similarity 100.0%; Pred. No. 2.5e-12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 342 VIEAPLACDRNEELAANYLLE 362

Db 340 VIEAPLACDRNEELAANYLLE 360

RESULT 3

ID Q8LA46 PRELIMINARY; PRT; 371 AA.
 AC Q8LA46;

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE DNA repair protein RAD23, putative.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 [1]
 RN SEQUENCE FROM N.A.
 RP Haas B.J., Volfovsky N., Town C.D., Troupkan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
 RT "Full-length messenger RNA sequences greatly improve genome
 RT annotation.";
 RL Genome Biol. 0:0-0(2002).
 [2]
 RN SEQUENCE FROM N.A.
 RP Brover V., Troupkan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RT "Full-length cDNA from Arabidopsis thaliana.";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY088037; AAM65503.1; -.
 DR InterPro; IPR004806; Rad23.
 DR InterPro; IPR006636; STI1.
 DR InterPro; IPR000449; UBA domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 2.
 DR SMART; SM00727; STI1; 1.
 DR SMART; SM00165; UBA; 2.
 DR SMART; SM00213; UBQ; 1.
 DR TIGRFAMs; TIGR00601; rad23; 1.
 DR PROSITE; PS0053; UBIQUITIN 2; 1.
 DR PROSITE; PS0053; UBIQUITIN 2; 1.
 SQ SEQUENCE 371 AA; 39747 MW; 081493086EA976E7 CRC64;
 Query Match 5.7%; Score 21; DB 10; Length 371;
 Best Local Similarity 100.0%; Pred. No. 2.5e-12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 342 VIEAFLACDRNEELAANYLLE 362
 DB 344 VIEAFLACDRNEELAANYLLE 364
 RESULT 4
 ID Q9SA20 PRELIMINARY; PRT; 113 AA.
 AC Q9SA20;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE F309.1 protein.
 GN F309.1
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=cv. Columbia;
 RA Vysotskaia V.S., Schwartz J., Yu G., Toriumi M., Lenz C., Liu S.,
 RA Lee J., Li J., Krenetskaia I., Liu A., Lueros J., Gonzalez A.,
 RA Altafi H., Araujo R., Chao Q., Conn L., Conway A.B., Dunn P.,
 RA Hansen N., Huizar L., Kim C., Palm C., Rowley D., Shinn P., Walker M.,
 RA Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;
 RT "The sequence of BAC F309 from Arabidopsis thaliana chromosome 1.";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=cv. Columbia;
 RC Theologis A.;
 RA Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC006341; AAD34676.1; -.
 DR HSSP; P54725; IDV0.
 DR InterPro; IPR000449; UBA_domain.
 DR Pfam; PF00627; UBA; 1.
 DR SMART; SM00165; UBA; 1.
 SQ SEQUENCE 113 AA; 12947 MW; 8BE7CB15C9E12380 CRC64;
 Query Match 4.1%; Score 15; DB 10; Length 113;
 Best Local Similarity 100.0%; Pred. No. 9.6e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 349 CORNEELAANYLLEH 363
 DB 93 CORNEELAANYLLEH 107
 RESULT 5
 ID Q94C35 PRELIMINARY; PRT; 378 AA.
 AC Q94C35;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical 40.1 kDa protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 [1]
 RN SEQUENCE FROM N.A.
 RP Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Shinn P.,
 RA Tracy S.E., Banh J., Bowser L., Carninci P., Chung M.K.,
 RA Goldsmith A.D., Hayashizaki Y., Iehida J., Jones T., Kamiya A.,
 RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
 RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Pham P.K., Quach H.L.,
 RA Sakano H., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C.,
 RA Toriumi M., Yamada K., Yu G., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cDNA clones.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY037181; AAK59766.1; -.
 DR InterPro; IPR004806; Rad23.
 DR InterPro; IPR006636; STI1.
 DR InterPro; IPR000449; UBA domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 2.
 DR Pfam; PF00240; ubiquitin; 1.
 DR SMART; SM00727; STI1; 1.
 DR SMART; SM00165; UBA; 2.
 DR SMART; SM00213; UBQ; 1.
 DR TIGRFAMs; TIGR00601; rad23; 1.
 DR PROSITE; PS0053; UBIQUITIN_2; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 378 AA; 40075 MW; E49B591527A3EC27 CRC64;
 Query Match 4.1%; Score 15; DB 10; Length 378;
 Best Local Similarity 100.0%; Pred. No. 3.1e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 174 NNPERAVEYLISGIP 188
 DB 177 NNPERAVEYLISGIP 191
 RESULT 6
 ID Q9FF16 PRELIMINARY; PRT; 378 AA.
 AC Q9FF16;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE DNA repair protein RAD23 homolog (Hypothetical protein)

(At5g38470/At5g38470).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 ON NCBI_TaxID=3702;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RX MEDLINE=97471969; PubMed=9330910;
 RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
 RA Miyajima N., Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
 RT features of the 1.6 Mb regions covered by twenty physically assigned
 RT pl. clones";
 RL DNA Res. 4:215-230 (1997).
 RN [2]
 RN SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cDNA clones";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Shinn P., Banh J.,
 RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
 RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis ORF clones";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RN SEQUENCE FROM N.A.
 RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
 RT "Full-length messenger RNA sequences greatly improve genome
 RT annotation.";
 RL Genome Biol. 0:0-0 (2002).
 RN [5]
 RN SEQUENCE FROM N.A.
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RT "Full-length cDNA from Arabidopsis thaliana";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB05248; BAB09359.1; -;
 DR EMBL; AY058196; AAL25609.1; -;
 DR EMBL; AY081835; AAL87405.1; -;
 DR EMBL; AY087564; AAM65106.1; -;
 DR HSSP; F54725; IDV0.
 DR InterPro; IPR004806; Rad23.
 DR InterPro; IPR004449; UBA domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 2.
 DR TIGRFAMs; TIGR00601; rad23; 1.
 DR PROSITE; PS50053; UBIQUITIN_2; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 378 AA; 40066 MW; DD9B59153EA3B6A9 CRC64;

Query Match 4.1%; Score 15; DB 10; Length 378;
 Best Local Similarity 100.0%; Pred. No. 3.1e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 174 NNPERAVELYSGIP 188

Db 177 NNPERAVELYSGIP 191

RESULT 7

003991 PRELIMINARY; PRT; 379 AA.
 ID O03991
 AC O03991; (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE RAD23 protein, isoform II.
 OS Daucus carota (Carrot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; campanulids; Apiales; Apiaceae; Daucus.
 ON NCBI_TaxID=4039;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=W001C;
 RX MEDLINE=98345997; PubMed=9681019;
 RA Sturm A., Leinhardt S.;
 RT "Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in
 RT yeast.";
 RL Plant J. 13:815-821 (1998).
 DR EMBL; Y12014; CAA72742.1; -;
 DR HSSP; P54725; IDV0
 DR InterPro; IPR004806; Rad23.
 DR InterPro; IPR006636; STIL.
 DR InterPro; IPR000449; UBA domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 2.
 DR SMART; SM00727; STIL; 1.
 DR SMART; SM00165; UBA; 2.
 DR SMART; SM00213; UBO; 1.
 DR TIGRFAMs; TIGR00601; rad23; 1.
 DR PROSITE; PS50053; UBIQUITIN_2; 1.
 SQ SEQUENCE 379 AA; 40530 MW; B26697B439CC5929 CRC64;

Query Match 4.1%; Score 15; DB 10; Length 379;
 Best Local Similarity 100.0%; Pred. No. 3.1e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 342 VIEAPLACDRNEELA 356

Db 352 VIEAPLACDRNEELA 366

RESULT 8

003990 PRELIMINARY; PRT; 382 AA.
 ID O03990
 AC O03990; (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE RAD23, isoform I.
 OS Daucus carota (Carrot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; campanulids; Apiales; Apiaceae; Daucus.
 ON NCBI_TaxID=4039;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=W001C;
 RX MEDLINE=98345997; PubMed=9681019;
 RA Sturm A., Leinhardt S.;
 RT "Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in
 RT yeast.";
 RL Plant J. 13:815-821 (1998).
 DR EMBL; Y12013; CAA72741.1; -;
 DR HSSP; P54725; IDV0.
 DR InterPro; IPR002965; P_rich_extensn.

DR InterPro; IPR004806; Rad23.
 DR InterPro; IPR006636; Still.
 DR InterPro; IPR000449; UBA domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 2.
 DR Pfam; PF00240; ubiquitin; 1.
 DR PRINTS; PR01217; PRICHEXTENS.
 DR SMART; SM00727; Still; 1.
 DR SMART; SM00165; UBA; 2.
 DR SMART; SM00213; UBQ; 1.
 DR TIGRFAMs; TIGR00601; rad23; 1.
 DR PROSITE; PS0053; UBIQUITIN 2; 1.
 SQ SEQUENCE 382 AA; 40344 MW; 27FF02A2402F3DC2 CRC64;

Query Match 4.1%; Score 15; DB 10; Length 382;
 Best Local Similarity 100.0%; Pred. No. 3.1e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 174 NNPERAVELYSGIP 188
 Db 187 NNPERAVELYSGIP 201
 |||||

RESULT 9

ID Q9S9L8 PRELIMINARY; PRT; 246 AA.
 AC Q9S9L8;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE T24D18.27 protein.
 GN T24D18.27.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids; II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 [1]

SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Liu S.X., Yu G., Sakano H., Jhaveri A., Lee J.M., Lenz C., Pham P.,
 RA Toriumi M., Chin C., Chlou J., Choi E., Chung M., Gonzalez A.,
 RA Hwang B., Koo T., Li J., Liu A., Vayberg M., Altafi H., Brooks S.,
 RA Buehler E., Chao Q., Conn L., Conway A., Hansen N., Johnson-Hopson C.,
 RA Khan S., Kim C., Lam B., Nguyen M., Palm C., Shinn P., Tambunga G.,
 RA Davis R.W., Ecker J.R., Federpiet N.A., Theologis A.;
 RT "The sequence of BAC T24D18 from Arabidopsis thaliana chromosome 1.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 [2]

SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Theologis A.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC010924; AAF18513.1; -.

DR HSSP; P02248; LUBI.
 DR InterPro; IPR000449; UBA domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 1.
 DR Pfam; PF00240; ubiquitin; 1.
 DR SMART; SM00165; UBA; 1.
 DR SMART; SM00213; UBQ; 1.
 DR PROSITE; PS0053; UBIQUITIN 2; 1.
 SQ SEQUENCE 246 AA; 25966 MW; 0867D1E7D72B6FFE CRC64;

Query Match 3.8%; Score 14; DB 10; Length 246;
 Best Local Similarity 100.0%; Pred. No. 2.1e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 167 RALRAAYNNPERAV 180
 Db 155 RALRAAYNNPERAV 168
 |||||

RESULT 10

Q9STA6 PRELIMINARY; PRT; 389 AA.
 AC Q9STA6;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE RAD23 protein.
 GN RAD23.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; Iamids; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 RN [1]
 SEQUENCE FROM N.A.
 RC STRAIN=cv. West Virginia 106; TISSUE=Fruit;
 RA Lemaire-Chanley M., Petit J., Raymond P., Chevallier C.;
 RT "Analysis of gene expression during early tomato fruit development by
 RT mRNA differential display.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ243875; CAB51544.1; -.
 DR HSSP; P54725; IDVO
 DR InterPro; IPR004806; Rad23.
 DR InterPro; IPR006636; Still.
 DR InterPro; IPR000449; UBA domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 2.
 DR Pfam; PF00240; ubiquitin; 1.
 DR SMART; SM00727; Still; 1.
 DR SMART; SM00165; UBA; 2.
 DR SMART; SM00213; UBQ; 1.
 DR TIGRFAMs; TIGR00601; rad23; 1.
 DR PROSITE; PS0053; UBIQUITIN 2; 1.
 SQ SEQUENCE 389 AA; 41508 MW; D2BAEDF0FE70778A CRC64;

Query Match 3.8%; Score 14; DB 10; Length 389;
 Best Local Similarity 100.0%; Pred. No. 3.3e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 167 RALRAAYNNPERAV 180
 Db 177 RALRAAYNNPERAV 190
 |||||

RESULT 11

ID Q40742 PRELIMINARY; PRT; 392 AA.
 AC Q40742;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE ORAD23.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoidae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 SEQUENCE FROM N.A.
 RC STRAIN=Nipponbare;
 RX MEDLINE=97369378; PubMed=9225866;
 RA Schultz T.F., Quatrano R.S.;
 RT "Characterization and expression of a rice RAD23 gene.";
 RL Plant Mol. Biol. 34:557-562(1997).
 DR EMBL; U63530; AAB65841.1; -.
 DR HSSP; P54725; IDVO.
 DR Gramene; Q40742; -.
 DR InterPro; IPR004806; Rad23.
 DR InterPro; IPR006636; Still.
 DR InterPro; IPR000449; UBA domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 2.

DR Pfam; PF00240; ubiquitous; 1.
 DR SMART; SM00727; STI1; 1.
 DR SMART; SM00165; UBA; 2.
 DR SMART; SM00213; URQ; 1.
 DR TIGRFAMs; TIGR00601; rad23; 1.
 DR PROSITE; PS50053; UBQUITIN_2; 1.
 SQ SEQUENCE 392 AA; 41754 MW; BD6E08574CC7CABC CRC64;

Query Match 3.5%; Score 13; DB 10; Length 392;
 Best Local Similarity 100.0%; Pred. No. 0.00034; Length 392;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 NNPERAVEYLVS 186
 |||||
 Db 187 NNPERAVEYLVS 199

RESULT 12

Q81MB7 Q81MB7 PRELIMINARY; PRT; 343 AA.
 AC Q81MB7
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE CG1836-PB.
 GN RAD23.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Scher S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Chalker S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers V.H., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt J., Andrews-Pfannkoch C., Baldwin D.,
 RA April J.F., Agbayani A., An H.J., Helt J., Helt J., Helt J., Helt J.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fessler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jalali M., Kalush F., Karp G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 RN [2]

RP SEQUENCE FROM N.A.
 RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dreenek D., Farfan D.,
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacleb J., Paragay V., Park S., Patel S., Pfeiffer B., Scheeler F.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of Drosophila melanogaster genome";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Bertram B., Carlson J.W., Celniker S.E.,
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of Drosophila melanogaster genome";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003844; AAN06526.1; --
 SQ SEQUENCE 343 AA; 37549 MW; 2DFC676B41852079 CRC64;

Query Match 3.0%; Score 11; DB 5; Length 343;
 Best Local Similarity 100.0%; Pred. No. 0.031; Length 343;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 YNNPERAVEYL 183
 |||||
 Db 113 YNNPERAVEYL 123

RESULT 13
 Q9XZE0 Q9XZE0 PRELIMINARY; PRT; 414 AA.
 AC Q9XZE0
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE DHR23.
 OS RAD23 OR DHR23 OR CG1836.
 GN Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nabrochkin E.N., Grischuk A.L., Soldatov A.V.;
 RT "Cloning and characterization of the Drosophila melanogaster homologue
 of the Saccharomyces cerevisiae gene RAD23";
 RL Genetika 35:0-0(1999).
 DR EMBL; AF136606; AAD33695.1; --
 DR HSSP; P54725; LDV0
 DR FlyBase; FBgn0026777; Rad23.
 DR InterPro; IPR004806; Rad23.
 DR InterPro; IPR006636; STI1.
 DR InterPro; IPR000449; UBA_domain.

DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 2.
 DR Pfam; PF00240; ubiquitin; 1.
 DR SMART; SM00727; STIL1; 1.
 DR SMART; SM00165; UBA; 2.
 DR SMART; SM00213; UBQ; 1.
 DR TIGRFAMs; TIGR00601; rad23; 1.
 DR PROSITE; PS0053; UBIQUITIN 2; 1.
 SQ SEQUENCE 414 AA; 45794 MW; 4C3FF94CAll16F7AB CRC64;

Query Match 3.0%; Score 11; DB 5; Length 414;
 Best Local Similarity 100.0%; Pred.No. 0.038; 0; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 YNNPRAVEYL 183
 Db 184 YNNPRAVEYL 194

RESULT 14

Q9V3W9
 ID Q9V3W9 PRELIMINARY; PRT; 414 AA.
 AC Q9V3W9
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE RAD23 protein.
 GN RAD23 OR CG1836.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke T.C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murthy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.

RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Brodsky M.H., Rubin G.M., Tsang G.;
 RT "Full length Drosophila melanogaster cDNA sequence.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB003844; AAF59352.1; -.
 DR EMBL; AF132147; AAD33594.1; -.
 DR HSSP; P54725; 1DV0.
 DR FlyBase; FBgn0026777; Rad23.
 DR InterPro; IPR004806; Rad23.
 DR InterPro; IPR006636; STIL.
 DR InterPro; IPR000449; UBA domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 2.
 DR Pfam; PF00240; ubiquitin; 1.
 DR SMART; SM00727; STIL; 1.
 DR SMART; SM00165; UBA; 2.
 DR SMART; SM00213; UBQ; 1.
 DR TIGRFAMs; TIGR00601; rad23; 1.
 DR PROSITE; PS0053; UBIQUITIN 2; 1.
 SQ SEQUENCE 414 AA; 45780 MW; 4C2E494CAll16F7AB CRC64;

Query Match 3.0%; Score 11; DB 5; Length 414;
 Best Local Similarity 100.0%; Pred.No. 0.038; 0; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 YNNPRAVEYL 183
 Db 184 YNNPRAVEYL 194

RESULT 15

Q9M887
 ID Q9M887 PRELIMINARY; PRT; 419 AA.
 AC Q9M887;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Putative RAD23 (AT3902540/F16B3_17).
 GN F16B3.17.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
 RA Rensing C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
 RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome III BAC F16B3 genomic sequence.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kim C.-J., Chen H., Cheuk R., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Gibson H.A., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
 RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
 RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
 RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Yamada K., Yu G., Yu S., Shinozaki K.,
 RA Davis R.W., Theologis A., Ecker J.R.;
 RT "Arabidopsis cDNA clones.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Kim C.-J., Chen H., Cheuk R., Shinn P., Banh J., Bowser L.,
 RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
 RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,

```

RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis ORF clones";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC021840; AAF32461.1; -.
DR EMBL; AY039562; AAK62617.1; -.
DR EMBL; AY113034; AAN47342.1; -.
DR HSSP; P54725; IDV0.
DR InterPro; IPR002965; P rich_extensn.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR000449; UBA domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR PRINTS; PR01217; PRICHEXTENS.
DR TIGRFRAMS; TIGR00601; rad23; 1.
DR PROSITE; PSS0053; UBIQUITIN_2; 1.
SQ SEQUENCE 419 AA; 44247 MW; CAA13BC4FFEB1E25 CRC64;

Query Match 3.0%; Score 11; DB 10; Length 419;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 VKTLKGTHPEI 15
| | | | |
Db 5 VKTLKGTHPEI 15

RESULT 16
Q9SCAB PRELIMINARY; PRT; 65 AA.
AC Q9SCAB;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Rad23 protein (Fragment).
GN RAD23.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. West virginia 106; TISSUE=Fruit;
RA Lemaire-Chamley M., Petit J., Causee M., Raymond P., Chevalier C.;
RT "Isolation and characterization of cDNAs expressed during early
RT development of tomato fruit by mRNA differential display.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ270958; CAB65692.1; -.
DR HSSP; P54725; IDV0.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF00627; UBA; 1.
DR SMART; SM00165; UBA; 1.
FT NON TER 1
SQ SEQUENCE 65 AA; 7233 MW; 0538945EE9030B4D CRC64;

Query Match 2.7%; Score 10; DB 10; Length 65;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 352 NEELAANYLL 361
| | | | |
Db 48 NEELAANYLL 57

RESULT 17
Q9RU37 PRELIMINARY; PRT; 225 AA.
AC Q9RU37
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

```

```

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE DNA-binding response regulator.
GN DR1558.
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Panghile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
CC -!- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
CC -!- REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS.
CC -!- SIMILARITY: BELONGS TO THE LUXR/UHPA FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL; AE001999; AAF11120.1; -.
DR HSSP; P10957; 1RNL.
DR TIGR; DR1558; -.
DR InterPro; IPR000792; HTH_LuxR.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF00196; GerE; 1.
DR Pfam; PF00072; response_reg; 1.
DR PRINTS; PR00038; HTHLUXR.
DR ProDom; PD000307; HTH_LuxR; 1.
DR ProDom; PD000039; Response_reg; 1.
DR SMART; SM00421; HTH_LuxR; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS0110; RESPONSE_REGULATORY; 1.
DR PROSITE; PS0110; SENSORY_TRANSDUCTION;
KW DNA-binding; Phosphorylation; Sensory transduction;
KW Transcription regulation; Complete proteome.
SQ SEQUENCE 225 AA; 24452 MW; 953B45742CB2A152 CRC64;

Query Match 2.4%; Score 9; DB 16; Length 225;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 STLEENKVN 65
| | | | |
Db 95 STLEENKVN 103

RESULT 18
Q9C552 PRELIMINARY; PRT; 753 AA.
AC Q9C552
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Putative RNA-binding domain (Hypothetical 84.8 kDa protein) (Putative
DE RNA-binding protein).
GN T10D10.21 OR T9N14.7.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopses.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altaji H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,

```

RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldbiyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Militischer J., Miranda M., Nguyen M., Nierman M.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vayberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.,
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana";
RL Nature 408:816-820(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Pham P.K., Banh J., Banno F., Dale J.M.,
RA Goldsmith A.D., Jiang P.X., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Yamamura Y., Yu G., Yu S., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koseena E., Lam B.,
RA Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Tracy S.E.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.,
RT "Full Length cDNA of gene T10D10.21 (GI:12325275).";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,
RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koseena E.,
RA Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,
RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.,
RT "Full Length cDNA of gene T10D10.21 (GI:12325275).";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC016529; AKG52582.1; -;
DR EMBL; AC067754; AKG51785.1; -;
DR EMBL; AY051073; AAK93750.1; -;
DR EMBL; AY035041; AAK95946.1; -;
DR InterPro; IPR001313; Pumilio/Puf.
DR Pfam; PF00806; PUF; 5.
DR SMART; SM00025; Pumilio; 5.
KW Hypothetical protein.
SQ SEQUENCE 753 AA; 84827 MW; 300CC905280C29CA CRC64;

Query Match 2.4%; Score 9; DB 10; Length 753;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 340 ARVIEAFLA 348
Db 542 ARVIEAFLA 550
|||||

RESULT 19
Q39682
ID Q39682 PRELIMINARY; PRT; 111 AA.
AC Q39682
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Glycine-rich protein (Fragment).
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Apiales; Apiaceae; Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=Danver Half-long; TISSUE=Somatic embryos at the globular stage;
RA Lin X., Hwang G.-J., Zimmerman J.L.;
RT "Isolation and characterization of a diverse set of genes from carrot
RT somatic embryos";
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U47097; AA01097.1; -;
FT NON TER 1
SQ SEQUENCE 111 AA; 13829 MW; 082789CCADF1C2A CRC64;

Query Match 2.2%; Score 8; DB 10; Length 111;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
Db 18 AGGGAGGG 25
|||||

RESULT 20
Q8CA15
ID Q8CA15 PRELIMINARY; PRT; 125 AA.
AC Q8CA15
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Hypothalamus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
DR EMBL; AK038724; BAC30110.1; -;
KW Hypothetical protein.
SQ SEQUENCE 125 AA; 13205 MW; 58AD04310B540101 CRC64;

Query Match 2.2%; Score 8; DB 11; Length 125;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
Db 82 AGGGAGGG 89
|||||

RESULT 21
Q9FOY9
ID Q9FOY9 PRELIMINARY; PRT; 139 AA.
AC Q9FOY9
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Avr9/Cf-9 rapidly elicited protein 75.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Petite Havana;
RA Durrant W.E., Rowland O., Piedras P., Hammond-Kosack K.E.,
RA Jones J.D.G.;
RT "cDNA expression profiling reveals rapid, resistance gene-dependent,
RT active oxygen-independent, gene induction during the plant defense

RT response.";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF211540; AAG43558.1; -.
 SQ SEQUENCE 139 AA; 14597 MW; 1D485BD90F811D9C CRC64;

Query Match 2.2%; Score 8; DB 10; Length 139;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGGG 241
 Db 87 AGGGAGGG 94
 |||||

RESULT 22

Q924J3 ID Q924J3 PRELIMINARY; PRT; 167 AA.
 AC Q924J3
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Interphotoreceptor binding protein (Fragment).
 GN IRBP.
 OS Tamas striatus (Eastern chipmunk).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Scuriidae; Scuriinae;
 OC Tamas.
 OC eurosids II; Brasciales; Arabidopsis.
 OX NCBI_TaxID=45474;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21240538; PubMed=11341810;
 RA DeBry R.W., Sagel R.M.;
 RT "Phylogeny of Rodentia (Mammalia) inferred from the nuclear-encoded
 RT gene IRBP";
 RL Mol. Phylogenet. Evol. 19:290-301(2001).
 DR EMBL; AF297282; AAK62261.1; -.
 DR InterPro; IPR003581; TSPC.
 DR Pfam; PF02692; IRBP; 1.
 KW Receptor.
 FT NON TER 1
 FT NON TER 167
 FT NON TER 167
 SQ SEQUENCE 167 AA; 18463 MW; 0ED660DB40D94819 CRC64;

Query Match 2.2%; Score 8; DB 11; Length 167;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 102 LEAPQAP 109
 Db 41 LEAPQAP 48
 |||||

RESULT 23

Q41191 ID Q41191 PRELIMINARY; PRT; 173 AA.
 AC Q41191
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Glycine-rich protein (Fragment).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brasciales; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93044485; PubMed=2152168;
 RA de Oliveira D.E., Seurinck J., Inze D., Van Montagu M., Botterman J.;
 RT "Differential expression of five Arabidopsis genes encoding glycine-
 RT rich proteins.";
 RL Plant Cell 2:427-436(1990).
 DR EMBL; S47414; AAB24077.1; -.

DR InterPro; IPR002952; Eggshell.
 DR InterPro; IPR000817; Prion.
 DR PRINTS; PRO1228; EGGSHLL.
 DR PRINTS; PRO0341; PRION.
 FT NON TER 1
 SQ SEQUENCE 173 AA; 13587 MW; D02259B4B38F80C3 CRC64;

Query Match 2.2%; Score 8; DB 10; Length 173;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGGG 241
 Db 84 AGGGAGGG 91
 |||||

RESULT 24

Q9LTP5 ID Q9LTP5 PRELIMINARY; PRT; 174 AA.
 AC Q9LTP5
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Genomic DNA, chromosome 3, Pl clone: MQC12 (At3g20470/MQC12_23).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brasciales; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RX MEDLINE=20277480; PubMed=10819329;
 RA Nakamura Y.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
 RT features of the regions of 4,504,864 bp covered by sixty Pl and TAC
 RT clones.";
 RL DNA Res. 7:131-135(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Koesema E., Chen H., Cheuk R., Kim C.J., Meyers M.C., Shinn P.,
 RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
 RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
 RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
 RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
 RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
 RT "Arabidopsis ORF clones.";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Gibson H.A., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
 RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
 RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
 RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Yamada K., Yu G., Yu S., Shinozaki K.,
 RA Davis R.W., Theologis A., Ecker J.R.;
 RT "Arabidopsis cDNA clones.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB024036; BAB02830.1; -.
 DR EMBL; AY058226; AAL15400.1; -.
 DR EMBL; AY045629; AAK73987.1; -.
 DR InterPro; IPR002952; Eggshell.
 DR InterPro; IPR000817; Prion.
 DR PRINTS; PRO1228; EGGSHLL.

DR PRINTS; PRO0341; PRION.
SQ SEQUENCE 174 AA; 13718 MW; 196DF768A69EA176 CRC64;

Query Match 2.2%; Score 8; DB 10; Length 174;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGGG 241
Db 85 AGGAGGG 92

RESULT 25

Q96V99 ID Q96V99 PRELIMINARY; PRT; 186 AA.
AC Q96V99;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Dihydrofolate reductase (SC 1.5.1.3) (Fragment).
OS Pneumocystis carinii f. sp. muris.
OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
OC Pneumocystis.
OC NCBI_TaxID=42066;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21537003; PubMed=11679931;
RA Ma L., Imachi H., Sukura A., Kovacs J.A.;
RT "Genetic Divergence of the Dihydrofolate Reductase and Dihydropteroate
RT Synthase Genes in Pneumocystis carinii from 7 Different Host
RT Species.";
RL J. Infect. Dis. 184:1358-1362 (2001).
DR EMBL; AF175561; AAL25644.1; -.
DR InterPro; IPR001796; DHFR.
DR Pfam; PF00186; Dihfolate_red; 1.
DR PRINTS; PRO0070; DHFR.
DR PROSITE; PS00075; DHFR; 1.
KW Oxidoreductase.
FT NON_TER 1 186
FT TER 186 186
SQ SEQUENCE 186 AA; 21402 MW; 1A181FDSAAAC8600 CRC64;

Query Match 2.2%; Score 8; DB 3; Length 186;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 KVNEDGFL 70
Db 178 KVNEDGFL 185

RESULT 26

Q9LGP0 ID Q9LGP0 PRELIMINARY; PRT; 202 AA.
AC Q9LGP0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE P0684C01.5 protein.
GN P0684C01.5.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OC NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
RT P0684C01.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002487; BAB07944.1; -.

DR Gramene; Q9LGP0; -.
SQ SEQUENCE 202 AA; 20969 MW; 44594A6CFE37B7F8 CRC64;

Query Match 2.2%; Score 8; DB 10; Length 202;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGGP 242
Db 24 GGGAGGGP 31

RESULT 27

Q8L418 ID Q8L418 PRELIMINARY; PRT; 208 AA.
AC Q8L418;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE P0551A11.6 protein (QJ1116 C07.6 protein).
GN P0551A11.6 OR QJ1116 C07.6
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OC NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
RT clone:P0551A11.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC
RT clone:QJ1116 C07.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003934; BAB92814.1; -.
DR EMBL; AP004253; BAC00689.1; -.
DR Gramene; Q8L418; -.
SQ SEQUENCE 208 AA; 20670 MW; C4D641DBA4D55B80 CRC64;

Query Match 2.2%; Score 8; DB 10; Length 208;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGGG 241
Db 71 AGGAGGG 78

RESULT 28

Q41187 ID Q41187 PRELIMINARY; PRT; 210 AA.
AC Q41187;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Glycine-rich protein (Fragment).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93044485; PubMed=2152168;
RA de Oliveira D.E., Seurinck J., Inze D., Van Montagu M., Botterman J.;
RT "Differential expression of five Arabidopsis genes encoding glycine-
RT rich proteins.";

```

RL Plant Cell 2:427-436(1990).
DR EMBL; S47405; AAB24073.1; -.
FT NON TER 1
SQ SEQUENCE 210 AA; 14445 MW; 03A5890F258204BC CRC64;

Query Match 2.2%; Score 8; DB 10; Length 210;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 241
Db 148 AGGAGGG 155

RESULT 29
Q39367
ID Q39367 PRELIMINARY; PRT; 220 AA.
AC Q39367;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glycine-rich protein (Fragment).
OS Brassica oleracea (Cauliflower).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3712;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=alaboglabra; TISSUE=Stigma;
RA Booker J.P.;
RL Thesis (1996), Biological Science, University of Durham.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=alaboglabra; TISSUE=Stigma;
RA Scutt C.P.;
RL Thesis (1990), Biological Science, University of Durham.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=alaboglabra; TISSUE=Stigma;
RA Croy R.D.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; 274892; CAA99171.1; -.
DR InterPro; IPR000817; Prion.
DR PRINTS; PR00341; PRION.
FT NON TER 1
SQ SEQUENCE 220 AA; 15368 MW; 7ED3ED15D717B519 CRC64;

Query Match 2.2%; Score 8; DB 10; Length 220;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 241
Db 16 AGGAGGG 23

RESULT 30
Q8C4G2
ID Q8C4G2 PRELIMINARY; PRT; 228 AA.
AC Q8C4G2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RA MEDLINE=22354683; PubMed=12466851;
RX

RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK082275; BAC38452.1; -.
KW Hypothetical protein.
SQ SEQUENCE 228 AA; 24880 MW; 0BAD53B878442E43 CRC64;

Query Match 2.2%; Score 8; DB 11; Length 228;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 241
Db 82 AGGAGGG 89

RESULT 31
Q8BPM6
ID Q8BPM6 PRELIMINARY; PRT; 228 AA.
AC Q8BPM6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Eye;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK053760; BAC35510.1; -.
KW Hypothetical protein.
SQ SEQUENCE 228 AA; 24908 MW; 5A7F826B78472F08 CRC64;

Query Match 2.2%; Score 8; DB 11; Length 228;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 241
Db 82 AGGAGGG 89

RESULT 32
Q8S7T7
ID Q8S7T7 PRELIMINARY; PRT; 307 AA.
AC Q8S7T7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 32.5 kDa protein (Fragment).
GN OSJNBAA009ip11.2.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R.; Yuan Q.; Ouyang S.; Liu J.; Moffat K.S.; Hill J.N.;
RA Gansberger K.; Brenner M.; Burgess S.; Hance M.; Shvartsbeyn M.;
RA Tstirin T.; Riggs F.; Hsiao J.; Zismann V.; Blunt S.; Pai G.;
RA VanAken S.E.; Utterback T.R.; Feldblyum T.V.; Kalb E.; Quackenbush J.;

```

RA Salzberg S.L., White O., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSUNBa0091P11 genomic sequence."
RL EMBL; AC073556; AAL84302.1; -.
DR Gramene; Q897T7; -.
DR InterPro; IPR006873; DUF620.
DR Pfam; PF04788; DUF620; 1.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 307 AA; 32503 MW; 8FA3FD1BED4BFBF9 CRC64;

Query Match 2.2%; Score 8; DB 10; Length 307;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGGG 241
DB 61 AGGAGGG 68
|||||

RESULT 33
Q9KYD8 PRELIMINARY; PRT; 340 AA.
AC Q9KYD8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Putative secreted transglycosylase.
GN SCO4132 OR SCD72A.18.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Brown S.P., Harris D.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Bentley S.D., Parkhill J., Barrell B.G., Rastam M.A.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapante D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rastam M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2)."
RL Nature 417:141-147(2002).
RL EMBL; AL939119; CAB92385.1; -.
DR InterPro; IPR000064; NLP_C60.
DR InterPro; IPR000189; SLT_domain.
DR Pfam; PF00877; NLP_C60; 1.
DR Pfam; PF01464; SLT; 1.

KW Complete proteome.
SQ SEQUENCE 340 AA; 35682 MW; 516F68207D31962A CRC64;

Query Match 2.2%; Score 8; DB 16; Length 340;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGGG 241
DB 40 AGGAGGG 47
|||||

RESULT 34
Q9Z2C8 PRELIMINARY; PRT; 360 AA.
AC Q9Z2C8;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-2003 (TrEMBLrel. 23, Last annotation update)
DE Y-box protein MSY2.
GN YBX2 OR MSY2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; Tissue=Testis;
RX MEDLINE=98455458; PubMed=9780336;
RA Gu W., Tekur S., Reinhold R., Spig J.J., Choi Y.C., Zheng J.Z.,
RA Murray M.T., Hecht N.B.;
RT "Mammalian male and female germ cells express a germ cell-specific Y-
box protein, MSY2."
RL Biol. Reprod. 59:1266-1274(1998).
CC - SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
DR EMBL; AF073954; AAC98673.1; -.
DR HSSP; P41016; 1C90.
DR MGD; MGI:1096372; Ybx2.
DR InterPro; IPR002059; Cold_shock.
DR Pfam; PF00313; CSD; 1.
DR PRINTS; PR00050; COLDSHOCK.
DR ProDom; PD000621; Cold_shock; 1.
DR SMART; SM00357; CSP; 1.
DR PROSITE; PS00352; COLD_SHOCK; 1.
SQ SEQUENCE 360 AA; 38270 MW; DFCAAE7F936731BF CRC64;

Query Match 2.2%; Score 8; DB 11; Length 360;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGGG 241
DB 40 AGGAGGG 47
|||||

RESULT 35
Q9GLB8 PRELIMINARY; PRT; 366 AA.
AC Q9GLB8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Interphotoreceptor retinoid-binding protein (Fragment).
GN IRBP.
OS Cephalorhynchus eutropia (Black dolphin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae;
OC Cephalorhynchus.
OX NCBI_TaxID=27608;
RN [1]
RP SEQUENCE FROM N.A.
RA Cassens I., Vicario S., Waddell V.G., Balchowsky H., Van Belle D.,
RA Ding W., Fan C., Mohan L., Simoes-Lopes P.C., Bastida R., Meyer A.,

```

RA Stanhope M.J., Milinkovitch M.C.;
RT "Independent adaptation to riverine habitats allowed survival of
RL ancient cetacean lineages.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF304076; AAC25970.1; -.
DR InterPro; IPR003581; TSPc.
DR Pfam; PF02692; IRBP; 1.
DR SMART; SM00245; TSPc; 1.
KW Receptor.
FT NON_TER 1 366
FT NON_TER 366 366
SQ SEQUENCE 366 AA; 39870 MW; 10A620359F3D1B7A CRC64;

Query Match 2.2%; Score 8; DB 6; Length 366;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 102 LEAPQAP 109
Db 1 LEAPQAP 8

RESULT 36
Q9P639
ID Q9P639 PRELIMINARY; PRT; 373 AA.
AC Q9P639;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Related to glycine-rich cell wall structural protein.
GN B21J21.090.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Align V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Meves H.W., Mannhaupt G.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL355929; CAB91331.2; -.
DR InterPro; IPR002952; Eggshell.
DR InterPro; IPR001185; MS_channel.
DR Pfam; PF01741; MscL; 1.
DR PRINTS; PR01228; EGGSHLL.
SQ SEQUENCE 373 AA; 36389 MW; 7546EBE7260ED192 CRC64;

Query Match 2.2%; Score 8; DB 3; Length 373;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGGG 241
Db 253 AGGAGGGG 260

RESULT 37
Q9R0J6
ID Q9R0J6 PRELIMINARY; PRT; 384 AA.
AC Q9R0J6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Interphotoreceptor retinoid binding protein (Fragment).
GN IRBP.
OS Apodemus speciosus (large Japanese field mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC Apodemus.

```

```

OX NCBI_TaxID=105296;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=HS240;
RX MEDLINE=20319711; PubMed=10862357;
RA Serizawa K., Suzuki H., Tsuchiya K.;
RT "A phylogenetic view on species radiation in Apodemus inferred from
RL variation of nuclear and mitochondrial genes.";
RL Biochem. Genet. 38:27-40(2000).
DR EMBL; AB032856; BAA85242.1; -.
DR InterPro; IPR003581; TSPc.
DR Pfam; PF02692; IRBP; 2.
DR SMART; SM00245; TSPc; 1.
KW Receptor.
FT NON_TER 1 384
FT NON_TER 384 384
SQ SEQUENCE 384 AA; 41642 MW; 8E4F40B292305814 CRC64;

Query Match 2.2%; Score 8; DB 11; Length 384;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 102 LEAPQAP 109
Db 26 LEAPQAP 33

RESULT 38
Q9R0H6
ID Q9R0H6 PRELIMINARY; PRT; 384 AA.
AC Q9R0H6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Interphotoreceptor retinoid binding protein (Fragment).
GN IRBP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Sprague-Dawley;
RX MEDLINE=20319711; PubMed=10862357;
RA Serizawa K., Suzuki H., Tsuchiya K.;
RT "A phylogenetic view on species radiation in Apodemus inferred from
RL variation of nuclear and mitochondrial genes.";
RL Biochem. Genet. 38:27-40(2000).
DR EMBL; AB033714; BAA85627.1; -.
DR InterPro; IPR003581; TSPc.
DR Pfam; PF02692; IRBP; 2.
DR SMART; SM00245; TSPc; 1.
KW Receptor.
FT NON_TER 1 384
FT NON_TER 384 384
SQ SEQUENCE 384 AA; 41790 MW; E4B6C02F02B50189 CRC64;

Query Match 2.2%; Score 8; DB 11; Length 384;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 102 LEAPQAP 109
Db 26 LEAPQAP 33

RESULT 39
Q9R0J5
ID Q9R0J5 PRELIMINARY; PRT; 384 AA.
AC Q9R0J5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

```

```

DE GN Interphotoreceptor retinoid binding protein (Fragment).
GN IRBP.
OS Apodemus peninsulae (Korean field mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC Apodemus.
OX NCBI_TaxID=105297;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS329/KT2772; PubMed=10862357;
RX MEDLINE=20319711; PubMed=10862357;
RA Serizawa K., Suzuki H., Tsuchiya K.;
RT "A phylogenetic view on species radiation in Apodemus inferred from
RT variation of nuclear and mitochondrial genes.";
RL Biochem. Genet. 38:27-40(2000).
DR EMBL; AB032857; BA85243.1; -.
DR InterPro; IPR003581; TSPC.
DR Pfam; PF02692; IRBP; 2.
DR SMART; SM00245; TSPC; 1.
KW Receptor.
FT NON_TER 1
FT NON_TER 384
FT SEQUENCE 384 AA; 41674 MW; F564540D3973EA2C CRC64;
SQ
Query Match 2.2%; Score 8; DB 11; Length 384;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 LEAPQAP 109
Db |||||
26 LEAPQAP 33

RESULT 40
Q9ROJ4 PRELIMINARY; PRT; 384 AA.
AC Q9ROJ4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE Interphotoreceptor retinoid binding protein (Fragment).
GN IRBP.
OS Apodemus agrarius (Eurasian field mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC Apodemus.
OX NCBI_TaxID=39030;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS1253/KT9271; PubMed=10862357;
RX MEDLINE=20319711; PubMed=10862357;
RA Serizawa K., Suzuki H., Tsuchiya K.;
RT "A phylogenetic view on species radiation in Apodemus inferred from
RT variation of nuclear and mitochondrial genes.";
RL Biochem. Genet. 38:27-40(2000).
DR EMBL; AB032858; BA85244.1; -.
DR InterPro; IPR003581; TSPC.
DR Pfam; PF02692; IRBP; 2.
DR SMART; SM00245; TSPC; 1.
KW Receptor.
FT NON_TER 1
FT NON_TER 384
FT SEQUENCE 384 AA; 41492 MW; BEBC5540C77E4D0F CRC64;
SQ
Query Match 2.2%; Score 8; DB 11; Length 384;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 LEAPQAP 109
Db |||||
26 LEAPQAP 33

```

```

RESULT 41
Q9ROH7 PRELIMINARY; PRT; 384 AA.
AC Q9ROH7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Interphotoreceptor retinoid binding protein (Fragment).
GN IRBP.
OS Tokudaia osimensis (Ryukyu spiny rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC Tokudaia.
OX NCBI_TaxID=73109;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS1162;
RX MEDLINE=20319711; PubMed=10862357;
RA Serizawa K., Suzuki H., Tsuchiya K.;
RT "A phylogenetic view on species radiation in Apodemus inferred from
RT variation of nuclear and mitochondrial genes.";
RL Biochem. Genet. 38:27-40(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HS1162;
RX MEDLINE=2029572; PubMed=10764531;
RA Suzuki H., Tsuchiya K., Takezaki N.;
RT "A molecular phylogenetic framework for the Ryukyu endemic rodents
RT Tokudaia osimensis and Diplothrix legata.";
RL Mol. Phylogenet. Evol. 15:15-24(2000).
DR EMBL; AB033712; BA85873.1; -.
DR InterPro; IPR003581; TSPC.
DR Pfam; PF02692; IRBP; 2.
DR SMART; SM00245; TSPC; 1.
KW Receptor.
FT NON_TER 1
FT NON_TER 384
FT SEQUENCE 384 AA; 41695 MW; AD57C8D5175F5475 CRC64;
SQ
Query Match 2.2%; Score 8; DB 11; Length 384;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 LEAPQAP 109
Db |||||
26 LEAPQAP 33

RESULT 42
Q8GVV7 PRELIMINARY; PRT; 387 AA.
AC Q8GVV7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE OJ1705_C03.15 protein.
GN OJ1705_C03.15.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, BAC
RT clone:OJ1705_C03.15."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003962; BAC45097.1; -.
SQ SEQUENCE 387 AA; 40202 MW; 8426D5490DFB09FB CRC64;

Query Match 2.2%; Score 8; DB 10; Length 387;

```

Best Local Similarity 100.0%; Pred. No. 39; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGGG 241
Db 40 AGGGAGGG 47

RESULT 43
Q9M3Y2
ID Q9M3Y2 PRELIMINARY; PRT; 390 AA.
AC Q9M3Y2;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Glycine-rich protein precursor.
GN GRP1.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21107239; PubMed=11161025;
RA Kingli C., Hauf G., Keller B.;
RT "Hydrophobic properties of the structural protein GRP1.8 in the cell
wall of protoxylem elements.";
RL Plant Physiol. 125:673-682(2001).
DR EMBL; AJ276509; CAB88804.1; -;
DR InterPro; IPR002952; Eggshell.
DR InterPro; IPR000817; Prion.
DR PRINTS; PR01228; EGGSHLL.
DR PRINTS; PR00341; PRION.
KW Signal.
FT SIGNAL. 1 28 POTENTIAL.
FT CHAIN 29 390 GLYCINE-RICH PROTEIN GRP1.
SQ SEQUENCE 390 AA; 28740 MW; 6EF5E2AA751A8B04 CRC64;

Query Match 2.2%; Score 8; DB 10; Length 390;
Best Local Similarity 100.0%; Pred. No. 39; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGGG 241
Db 83 AGGGAGGG 90

RESULT 44
Q64217
ID Q64217 PRELIMINARY; PRT; 395 AA.
AC Q64217;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Interphotoreceptor retinoid binding protein (Fragment).
OS Spalax zemni (Podolsk mole rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Spalacinae;
OC Spalax.
OX NCBI_TaxID=42150;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96304325; PubMed=8660440;
RA Stanhope M.J., Smith M.R., Waddell V.G., Porter C.A., Shivji M.S.,
RA Goodman M.;
RT "Mammalian evolution and the interphotoreceptor retinoid binding
protein (IRBP) gene: convincing evidence for several superordinal
clades.";
RL J. Mol. Evol. 43:83-92(1996).
DR EMBL; U48589; AAB18667.1; -;
DR InterPro; IPR003581; TSPC.
DR Pfam; PF02692; IRBP; 1.

DR SMART; SM00245; TSPC; 1.
KW Receptor.
FT NON TER 1
FT NON TER 395
SQ SEQUENCE 395 AA; 42331 MW; F5391D0D4DD1D266 CRC64;

Query Match 2.2%; Score 8; DB 11; Length 395;
Best Local Similarity 100.0%; Pred. No. 39; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 102 LEAPQQAP 109
Db 53 LEAPQQAP 60

RESULT 45
Q8GTL0
ID Q8GTL0 PRELIMINARY; PRT; 400 AA.
AC Q8GTL0;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Putative glycine-rich cell wall protein precursor.
GN QJ1165_F02.18.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
clone:QJ1165_F02.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003816; BAC21356.1; -;
SQ SEQUENCE 400 AA; 28516 MW; 84B816F150070E60 CRC64;

Query Match 2.2%; Score 8; DB 10; Length 400;
Best Local Similarity 100.0%; Pred. No. 40; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGGG 241
Db 161 AGGGAGGG 168

RESULT 46
Q9C518
ID Q9C518 PRELIMINARY; PRT; 401 AA.
AC Q9C518;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Auxin-induced basic helix-loop-helix transcription factor,
putative.
GN T18124.2 OR T15W6.11 OR ATLGS8100.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,

```

RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,
RA Militecher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Sallberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carrinci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka Y., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC079131; AAG50759.1; -.
DR EMBL; AC079604; AAG50894.1; -.
DR EMBL; AY081344; AAL91233.1; -.
DR InterPro; IPR005333; TCP.
DR Pfam; PF03634; TCP; 1.
SQ SEQUENCE 401 AA; 42471 MW; 9F58817CDB11874C CRC64;

Query Match      2.2%; Score 8; DB 10; Length 401;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
DB 294 AGGGAGGG 301

RESULT 47
Q8H010
ID Q8H010 PRELIMINARY; PRT; 420 AA.
AC Q8H010;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE Hypothetical protein.
GN O1081D05:6.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Wing R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Currie J.,
RA Collura K.;
RT "Rice Genomic Sequence.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; ACL34229; AAO06972.1; -.
KW Hypothetical protein.
SQ SEQUENCE 420 AA; 44730 MW; 344B8622E20ABFCE CRC64;

Query Match      2.2%; Score 8; DB 10; Length 420;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
DB 81 AGGGAGGG 88

RESULT 48
Q8K1Q2
ID Q8K1Q2 PRELIMINARY; PRT; 423 AA.
AC Q8K1Q2;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Interphotoreceptor retinoid binding protein (Fragment).
GN IRBP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Huchon D., Madsen O., Sibbald M.J.J.B., Ament K., Stanhope M.J.,
RA Catzeflis F., de Jong W.W., Douzery E.J.P.;
RT "Rodent phylogeny and a timescale for the evolution of Glires:
RT evidence from an extensive taxon sampling using three nuclear genes.";
RL Mol. Biol. Evol. 0:0-0 (2002).
DR EMBL; AJ429134; CAD22102.1; -.
DR InterPro; IPR003581; TSPC.
DR Pfam; PF02692; IRBP; 2.
DR SMART; SM00245; TSPC; 1.
KW Receptor.
FT NON_TER 1
FT NON_TER 423
FT NON_TER 424
SQ SEQUENCE 423 AA; 46154 MW; DB96D43C18CA73E5 CRC64;

Query Match      2.2%; Score 8; DB 11; Length 423;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 LEAQQOAP 109
DB 54 LEAQQOAP 61

RESULT 49
Q8K1P1
ID Q8K1P1 PRELIMINARY; PRT; 424 AA.
AC Q8K1P1;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Interphotoreceptor retinoid binding protein (Fragment).
GN IRBP.
OS Tachyoryctes sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Tachyoryctes.
OX NCBI_TaxID=123371;
RN [1]
RP SEQUENCE FROM N.A.
RA Huchon D., Madsen O., Sibbald M.J.J.B., Ament K., Stanhope M.,
RA Catzeflis F., de Jong W.W., Douzery E.J.P.;
RT "Rodent phylogeny and a timescale for the evolution of Glires:
RT evidence from an extensive taxon sampling using three nuclear genes.";
RL Mol. Biol. Evol. 0:0-0 (2002).
DR EMBL; AJ427231; CAD20269.1; -.
DR InterPro; IPR003581; TSPC.
DR Pfam; PF02692; IRBP; 2.
DR SMART; SM00245; TSPC; 1.
KW Receptor.
FT NON_TER 1
FT NON_TER 424
FT NON_TER 424
SQ SEQUENCE 424 AA; 45978 MW; 21EF2B620EB439F2 CRC64;

Query Match      2.2%; Score 8; DB 11; Length 424;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 LEAQQOAP 109
DB 54 LEAQQOAP 61

```

Db 55 LEAPQAP 62

RESULT 50

Q9GR85 PRELIMINARY; PRT; 427 AA.
 AC Q9GR85;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE HTLC2 protein.
 GN HTLC2.
 OS Halocynthia roretzi (Sea squirt).
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
 OC Stolidobranchia; Pyridae; Halocynthia.
 OX NCBI_taxID=7729;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Takahashi H., Satoh N.;
 RT "Trunk lateral cell-specific genes of ascidian Halocynthia roretzi."; Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB053355; BAB20902.1; -.
 DR HSPF; P19339; ISXL.
 DR InterPro; IPR002343; Hud_Sx1_RNA.
 DR Pfam; PF00076; xrm; 2.
 DR PRINTS; PR00961; HUSDXLRNA.
 DR SMART; SM00360; RRM; 2.
 DR PROSITE; PS0102; RRM; 2.
 DR PROSITE; PS00030; RRM_RNP_1; 2.
 SQ SEQUENCE 427 AA; 46483 MW; 28017858775EB032 CRC64;

Query Match 2.2%; Score 8; DB 5; Length 427;

Best Local Similarity 100.0%; Pred. No. 42;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 TSSSQHSN 92

Db 37 TSSSQHSN 44

RESULT 51

Q9NG13 PRELIMINARY; PRT; 438 AA.
 AC Q9NG13;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Intermediate filament protein C2 (Fragment).
 GN C2.
 OS Branchiostoma floridae (Florida lancelet) (Amphioxus).
 OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
 OC Branchiostoma.
 OX NCBI_taxID=7739;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20174564; PubMed=10711422;
 RA Karabinos A., Riemer D., Panopoulou G., Lehrach H., Weber K.;
 RT "Characterisation and tissue-specific expression of the two keratin subfamilies of intermediate filament proteins in the cephalochordate Branchiostoma."; Eur. J. Cell Biol. 79:17-26(2000).
 RL EMBL; AJ245429; CAB75938.1; -.
 DR InterPro; IPR001664; IF.
 DR InterPro; IPR002957; Keratin_I.
 DR Pfam; PF00038; filament; 1.
 DR PRINTS; PR01248; TYPE1KERATIN.
 DR PROSITE; PS00226; IF; 1.
 FT NON_TER 1
 FT NON_TER 438
 SQ SEQUENCE 438 AA; 47162 MW; 5958FE10F6ED979B CRC64;

Query Match 2.2%; Score 8; DB 5; Length 438;

Best Local Similarity 100.0%; Pred. No. 43;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 241

Db 429 AGGAGGG 436

RESULT 52

Q8PJY7 PRELIMINARY; PRT; 458 AA.
 AC Q8PJY7;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE ATP-dependent RNA helicase.
 GN DBPA OR XAC2390.
 OS Xanthomonas axonopodis (pv. citri).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_taxID=92829;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=306 / ATCC 13902 / XV 101;
 RX MEDLINE=2202145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.B.A.,
 RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Clapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., Oliveira V.R.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spinola L.A.F., Takita M.A., Tamura R.B., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.P.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing host specificities."; Nature 417:459-463(2002).
 RL Nature 417:459-463(2002).
 DR EMBL; AB011876; AM37242.1; -.
 DR InterPro; IPR005580; DbpA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR000629; DEAD_box.
 DR InterPro; IPR001650; Helicase_C.
 DR Pfam; PF03880; DbpA; 1.
 DR Pfam; PF00270; DEAD; 1.
 DR Pfam; PF00271; helicase_C; 1.
 DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
 KW Helicase; Complete proteome.
 SQ SEQUENCE 458 AA; 49269 MW; 44A78F9B9BB06EC CRC64;

Query Match 2.2%; Score 8; DB 16; Length 458;

Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 210 TGEAGLSG 217

Db 407 TGEAGLSG 414

RESULT 53

Q8P8G0 PRELIMINARY; PRT; 458 AA.
 AC Q8P8G0;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE ATP-dependent RNA helicase.
 GN DBPA OR XCC2282.


```

OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC NCBI_TaxID=340;
OX [1]_SEQUENCE FROM N.A.
RN
RP STRAIN=ATCC 33913 / NCPPB 528;
RC MEDLINE=2022145; PubMed=11202417;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RX Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chamberg P., Clapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.P., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL; AEO12335; AAM41561.1; --
DR InterPro; IPR005580; DbpA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR000629; DEAD_box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF03880; DbpA; 1.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; helicase_C; 1.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW Helicase; Complete proteome.
SQ SEQUENCE 458 AA; 49312 MW; CEE5277ACB98C438 CRC64;

Query Match 2.2%; Score 8; DB 16; Length 458;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 210 TGEAGLSG 217
Db 407 TGEAGLSG 414

RESULT 54
Q9UNW9 PRELIMINARY; PRT; 492 AA.
AC Q9UNW9; Q9UEAL;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE RNA-binding protein Nova-2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99007301; PubMed=9789075;
RA Yang Y.Y., Yin G.L., Darnell R.B.;
RT "The neuronal RNA-binding protein Nova-2 is implicated as the
RT autoantigen targeted in FOMIA patients with dementia.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:13254-13259(1998).
RN [2]
RP SEQUENCE OF 29-492 FROM N.A.
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K.J., Gordon L., Dias J., Ramirez M., Stilwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Barnes J.,
RA Danganan L., Brler A., Christensen M., Georgescu A., Avila J., Liu S.,
RA Attix C., Andreise T., Trankheim M., Amico-Keller G., Coefield J.,

```

```

RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,
RA Arellano A., Brower A., Sanders C., Ow D., Nolan M., Trong S.,
RA Kobayashi A., Olsen A.S., Carriano A.V.;
RT "Sequence analysis of a 1.9 Mb region in 19q13.2 between APOE and
RT D19S412.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 3 KH DOMAINS.
DR EMBL; AF038898; AAC72355.1; --
DR EMBL; AC006540; AAD13116.1; --
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH_type_1.
DR Pfam; PF00013; KH; 3.
DR SMART; SM00322; KH; 3.
DR PROSITE; PS00084; KH TYPE 1; 3.
SQ SEQUENCE 492 AA; 49008 MW; 41B63EAF6899256B CRC64;

Query Match 2.2%; Score 8; DB 4; Length 492;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 234 AGGGAGGG 241
Db 370 AGGGAGGG 377

RESULT 55
Q35392 PRELIMINARY; PRT; 492 AA.
AC Q35392;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Forkhead 2.
GN FOXD2 OR MF2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mesenchyme;
RX MEDLINE=98168839; PubMed=9510020;
RA Wu S.C.-Y., Grindley J., Winnier G.E., Hargett L., Hogan B.L.M.;
RT "Mouse Mesenchyme forkhead 2 (Mf2): expression, DNA binding and
RT induction by sonic hedgehog during somitogenesis.";
RL Mech. Dev. 70:3-13(1998).
DR EMBL; AF023915; AAB81275.1; --
DR HSP; Q63245; 2HPH.
DR TRANSFAC; T02492; --
DR MGD; MGI:1347471; Foxd2.
DR InterPro; IPR001766; TF_Fork_head.
DR Pfam; PF00250; Fork_head; 1.
DR PRINTS; PR00053; FORKHEAD.
DR ProDom; PD000425; TF_Fork_head; 1.
DR SMART; SM00339; FH; 1.
DR PROSITE; PS00657; FORK_HEAD_1; 1.
DR PROSITE; PS00658; FORK_HEAD_2; 1.
DR PROSITE; PS00039; FORK_HEAD_3; 1.
SQ SEQUENCE 492 AA; 48936 MW; 7F82440F4C435702 CRC64;

Query Match 2.2%; Score 8; DB 11; Length 492;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 234 AGGGAGGG 241
Db 392 AGGGAGGG 399

RESULT 56
Q43267 PRELIMINARY; PRT; 498 AA.
AC Q43267;

```

```

DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE ASTROCYTIC NOVA1-like RNA-binding protein (NEUROONCOLOGIC ventral
DE antigen 3) (Fragment).
GN ANOVA OR NOVA3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=20197319; PubMed=10735272;
RA Ueki K., Ramaswamy S., Billings S.J., Mohrenweiser H.W., Louis D.N.;
RT "ANOVA, a putative astrocytic RNA binding protein gene that maps to
RT chromosome 19q13.3." (1997).
RL Neurogenetics 1:31-36(1997).
CC -1- FUNCTION: MAY REGULATE RNA SPLICING OR METABOLISM IN A SPECIFIC
CC SUBSET OF DEVELOPING NEURONS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: BRAIN. EXPRESSION RESTRICTED TO ASTROCYTES.
CC -1- DISEASE: MAY BE A TARGET ANTIGEN IN ONE OF THE UNDEFINED HUMAN
CC PARANEOPLASTIC SYNDROMES.
DR EMBL; U70477; AAB8661.1; -.
DR Genew; HGNC:7887; NOVA2.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH_type_1.
DR Pfam; PF00013; KH; 3.
DR SMART; SM00322; KH; 3.
DR PROSITE; PS00084; KH_TYPE_1; 3.
KW Nuclear protein; RNA-binding; Repeat; Antigen.
FT NON_TER 1
FT DOMAIN 1 42 76 KH.
FT DOMAIN 140 174 KH.
FT DOMAIN 246 255 POLY-ALA.
FT DOMAIN 325 330 POLY-ALA.
FT DOMAIN 350 356 POLY-PRO.
FT DOMAIN 363 368 POLY-ALA.
FT DOMAIN 375 385 POLY-GLY.
FT DOMAIN 389 397 POLY-ALA.
FT DOMAIN 416 446 KH.
SQ SEQUENCE 498 AA; 49721 MW; C4B54196FDB6BF78 CRC64;

Query Match 2.2%; Score 8; DB 4; Length 498;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGGG 241
Db 376 AGGGAGGG 383

RESULT 57
Q816C7 ID Q816C7 PRELIMINARY; PRT; 544 AA.
AC Q816C7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG3588 (Fragment).
GN CG3588.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Zimbabwe/Hare19;
RX MEDLINE=22247773; PubMed=12351680;
RA Harr B., Kauer M., Schlotter C.;
RT "Hitchhiking mapping: A population-based fine-mapping strategy for
RT adaptive mutations in Drosophila melanogaster.";

DT adaptive mutations in Drosophila melanogaster.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12949-12954 (2002).
FT EMBL; AF495952; AAN09997.1; -.
FT NON_TER 1
FT NON_TER 544
SQ SEQUENCE 544 AA; 55701 MW; 640FA3F264821E75 CRC64;

Query Match 2.2%; Score 8; DB 5; Length 544;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 233 NAGGGAGG 240
Db 516 NAGGGAGG 523

RESULT 58
Q816C6 ID Q816C6 PRELIMINARY; PRT; 544 AA.
AC Q816C6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG3588 (Fragment).
GN CG3588.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Zimbabwe/Hare20;
RX MEDLINE=22247773; PubMed=12351680;
RA Harr B., Kauer M., Schlotter C.;
RT "Hitchhiking mapping: A population-based fine-mapping strategy for
RT adaptive mutations in Drosophila melanogaster.";

DT adaptive mutations in Drosophila melanogaster.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12949-12954 (2002).
FT EMBL; AF495953; AAN09998.1; -.
FT NON_TER 1
FT NON_TER 544
SQ SEQUENCE 544 AA; 55771 MW; 9FDA1909E06A17AB CRC64;

Query Match 2.2%; Score 8; DB 5; Length 544;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 233 NAGGGAGG 240
Db 516 NAGGGAGG 523

RESULT 59
Q816C5 ID Q816C5 PRELIMINARY; PRT; 544 AA.
AC Q816C5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG3588 (Fragment).
GN CG3588.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kenia11;
RX MEDLINE=22247773; PubMed=12351680;
RA Harr B., Kauer M., Schlotter C.;
RT "Hitchhiking mapping: A population-based fine-mapping strategy for
RT adaptive mutations in Drosophila melanogaster.";

```

RL Proc. Natl. Acad. Sci. U.S.A. 99:12949-12954 (2002).
 DR EMBL; AF495954; AAN09994.1; -.
 FT NON TER 1 1
 FT NON TER 544 544
 SQ SEQUENCE 544 AA; 55757 MW; 774BD4F723C22386 CRC64;

Query Match 2.2%; Score 8; DB 5; Length 544;
 Best Local Similarity 100.0%; Pred. No. 54;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 NAGGGAGG 240
 |||||
 Db 516 NAGGGAGG 523

RESULT 60

ID Q816C4 PRELIMINARY; PRT; 544 AA.
 AC Q816C4;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE CG3588 (Fragment).
 GN CG3588.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Kenial14;
 RX MEDLINE=22247773; PubMed=12351680;
 RA Harr B., Kauer M., Schloetterer C.;
 RT "Hitchhiking mapping: A population-based fine-mapping strategy for
 RT adaptive mutations in Drosophila melanogaster.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:12949-12954 (2002).
 DR EMBL; AF495955; AAN10000.1; -.
 FT NON TER 1 1
 FT NON TER 544 544
 SQ SEQUENCE 544 AA; 55691 MW; 619CB3335141D9389 CRC64;

Query Match 2.2%; Score 8; DB 5; Length 544;
 Best Local Similarity 100.0%; Pred. No. 54;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 NAGGGAGG 240
 |||||
 Db 516 NAGGGAGG 523

RESULT 61

ID Q816C3 PRELIMINARY; PRT; 544 AA.
 AC Q816C3;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE CG3588 (Fragment).
 GN CG3588.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Kenial17;
 RX MEDLINE=22247773; PubMed=12351680;
 RA Harr B., Kauer M., Schloetterer C.;
 RT "Hitchhiking mapping: A population-based fine-mapping strategy for
 RT adaptive mutations in Drosophila melanogaster.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:12949-12954 (2002).

DR EMBL; AF495956; AAN10001.1; -.
 FT NON TER 1 1
 FT NON TER 544 544
 SQ SEQUENCE 544 AA; 55757 MW; 4ADF86A1215BEEA3 CRC64;

Query Match 2.2%; Score 8; DB 5; Length 544;
 Best Local Similarity 100.0%; Pred. No. 54;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 NAGGGAGG 240
 |||||
 Db 516 NAGGGAGG 523

RESULT 62

ID Q816B7 PRELIMINARY; PRT; 544 AA.
 AC Q816B7;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE CG3588 (Fragment).
 GN CG3588.
 OS Drosophila simulans (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7240;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=New Caledonia;
 RX MEDLINE=22247773; PubMed=12351680;
 RA Harr B., Kauer M., Schloetterer C.;
 RT "Hitchhiking mapping: A population-based fine-mapping strategy for
 RT adaptive mutations in Drosophila melanogaster.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:12949-12954 (2002).
 DR EMBL; AF495972; AAN10017.1; -.
 FT NON TER 1 1
 FT NON TER 544 544
 SQ SEQUENCE 544 AA; 55449 MW; 9CF8E3EB9EB7A755 CRC64;

Query Match 2.2%; Score 8; DB 5; Length 544;
 Best Local Similarity 100.0%; Pred. No. 54;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 NAGGGAGG 240
 |||||
 Db 516 NAGGGAGG 523

RESULT 63

ID Q81072 PRELIMINARY; PRT; 544 AA.
 AC Q81072;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE CG3588 (Fragment).
 GN CG3588.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Zimbabwe/Harare12, and Zimbabwe/Harare18;
 RX MEDLINE=22247773; PubMed=12351680;
 RA Harr B., Kauer M., Schloetterer C.;
 RT "Hitchhiking mapping: A population-based fine-mapping strategy for
 RT adaptive mutations in Drosophila melanogaster.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:12949-12954 (2002).
 DR EMBL; AF495949; AAN09994.1; -.
 FT NON TER 1 1
 FT NON TER 544 544
 SQ SEQUENCE 544 AA; 55449 MW; 9CF8E3EB9EB7A755 CRC64;

Query Match 2.2%; Score 8; DB 5; Length 544;
 Best Local Similarity 100.0%; Pred. No. 54;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 NAGGGAGG 240
 |||||
 Db 516 NAGGGAGG 523

RESULT 63

ID Q81072 PRELIMINARY; PRT; 544 AA.
 AC Q81072;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE CG3588 (Fragment).
 GN CG3588.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Zimbabwe/Harare12, and Zimbabwe/Harare18;
 RX MEDLINE=22247773; PubMed=12351680;
 RA Harr B., Kauer M., Schloetterer C.;
 RT "Hitchhiking mapping: A population-based fine-mapping strategy for
 RT adaptive mutations in Drosophila melanogaster.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:12949-12954 (2002).
 DR EMBL; AF495949; AAN09994.1; -.
 FT NON TER 1 1
 FT NON TER 544 544
 SQ SEQUENCE 544 AA; 55449 MW; 9CF8E3EB9EB7A755 CRC64;

```
DR EMBL: AF495951; AAN09996.1; -.
FT NON_TER 1 544
FT NON_TER 544
SQ SEQUENCE 544 AA; 55729 MW; DEDALCODE16ALED2 CRC64;

Query Match 2.2%; Score 8; DB 5; Length 544;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 233 NAGGAGG 240
Db 516 NAGGAGG 523

RESULT 64
Q22809 Q22809 PRELIMINARY; PRT; 548 AA.
AC Q22809;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 61.8 kDa protein.
GN T26C11.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderiinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Martin J.;
RT "The sequence of C. elegans cosmid T26C11."
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: U41017; AAC48213.1; -.
DR WormPep; T26C11.4; CE05023.
KW Hypothetical protein.
SQ SEQUENCE 548 AA; 61841 MW; 0BF798A10B148E03 CRC64;

Query Match 2.2%; Score 8; DB 5; Length 548;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 228 PQGASNAG 235
Db 195 PQGASNAG 202

RESULT 65
Q81GB4 Q81GB4 PRELIMINARY; PRT; 550 AA.
AC Q81GB4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE RH70633p.
GN CG11711.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
```

```
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=y;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dreesnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BT001864; AAN71631.1; -.
SQ SEQUENCE 550 AA; 57774 MW; 36A7B8CAD2B5E412 CRC64;

Query Match 2.2%; Score 8; DB 5; Length 550;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGGG 241
Db 95 AGGGAGGG 102

RESULT 66
Q9VTJ7 Q9VTJ7 PRELIMINARY; PRT; 566 AA.
AC Q9VTJ7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE CG11711 protein.
GN CG11711 OR CG11712.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler K., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., M.G.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shie B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
```

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
DR EMBL; AF003544; AAF50051.1; -.
DR FlyBase; FBgn0036172; CG11711.
DR InterPro; IPR005301; Mobl_phocoin.
DR Pfam; PF03637; Mobl_phocoin; 1.
SQ SEQUENCE 566 AA; 58907 MW; 1D3F57A2659DEC2C CRC64;

Query Match 2.2%; Score 8; DB 5; Length 566;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGGG 241
Db 95 AGGGAGGG 102
|||||

RESULT 67
Q8LRM4 PRELIMINARY; PRT; 579 AA.
AC Q8LRM4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative amt protein.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-9;
RA Yao T., Feild E., King N., Soupen E., Kim K.-S., Kustu S.;
RT "Chlamydomonas reinhardtii amt gene.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF509497; AAM43911.1; -.
DR InterPro; IPR001905; Ammonium_transpt.
DR Pfam; PF00909; Ammonium_transp; 1.
DR TIGRFAMs; TIGR00836; amt; 1.
DR PROSITE; PS01219; AMMONIUM_TRANS; 1.
SQ SEQUENCE 579 AA; 60697 MW; 170C49CF09FB5C2D CRC64;

Query Match 2.2%; Score 8; DB 10; Length 579;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGGG 241
Db 535 AGGGAGGG 542
|||||

RESULT 68
Q8LRM5 PRELIMINARY; PRT; 579 AA.
AC Q8LRM5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative amt protein.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CC-125;
RA Yao T., Feild E., King N., Soupen E., Kim K.-S., Kustu S.;
RT "Chlamydomonas reinhardtii amt gene.";

RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF509496; AAM43910.1; -.
DR InterPro; IPR001905; Ammonium_transpt.
DR Pfam; PF00909; Ammonium_transp; 1.
DR TIGRFAMs; TIGR00836; amt; 1.
DR PROSITE; PS01219; AMMONIUM_TRANS; 1.
SQ SEQUENCE 579 AA; 60683 MW; B20C4CCA09FB5C2B CRC64;

Query Match 2.2%; Score 8; DB 10; Length 579;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGGG 241
Db 535 AGGGAGGG 542
|||||

RESULT 69
Q9FCJ3 PRELIMINARY; PRT; 584 AA.
AC Q9FCJ3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein SC05195.
GN SC05195 OR 2SC3B6.19.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cardeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb *Streptomyces coelicolor* A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete *Streptomyces
coelicolor* A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939122; CAC01325.1; -.
DR InterPro; IPR002838; DUF124.
DR InterPro; IPR002965; P_rich_extensn.
DR InterPro; IPR003325; TerD.
DR Pfam; PF02342; TerD; 1.
DR PRINTS; PR01217; PRICHEXTENS.
DR ProDom; PD013634; DUF124; 1.
KW Hypothetical protein; Complete proteome.

```

SQ SEQUENCE 584 AA; 60678 MW; 5DD7851F4DB7A76F CRC64;
Query Match 2.2%; Score 8; DB 16; Length 584;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 AQQAPQP 111
DB 181 AQQAPQP 188

RESULT 70
O43981 PRELIMINARY; PRT; 712 AA.
AC O43981;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Microneme protein Etmic-1 precursor.
GN MIC-1.
OS Eimeria tenella.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;
OC Eimeria.
OX NCBI_TaxID=5802;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Houghton;
RX MEDLINE=92131064; PubMed=1775171;
RA Tomley F.M., Clarke L.E., Kawazoe U., Dijkema R., Kok J.J.;
RT "Sequence of the gene encoding an immunodominant microneme protein of
Eimeria tenella.";
RL Mol. Biochem. Parasitol. 49:277-288(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Houghton;
RX MEDLINE=99094493;
RA Kelleher M., Tomley F.M.;
RT "Transient expression of beta-galactosidase in differentiating
sporozoites of Eimeria tenella.";
RL Mol. Biochem. Parasitol. 97:21-31(1998).
DR EMBL; AF032905; AAD03350.1; -.
DR HSSP; P17301; IAOX.
DR InterPro; IPR000884; TSPI.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF00090; tsp_1; 6.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00209; TSPI; 6.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS50092; TSPI; 5.
DR PROSITE; PS50234; VWFA; 1.
KW Signal.
FT SIGNAL.
FT CHAIN 37 712 MICRONEME PROTEIN ETMIC-1.
SQ SEQUENCE 712 AA; 74777 MW; 15B8F3C190870F73 CRC64;

Query Match 2.2%; Score 8; DB 5; Length 712;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 NAGGGAGG 240
DB 612 NAGGGAGG 619

RESULT 71
Q9AWU2 PRELIMINARY; PRT; 731 AA.
AC Q9AWU2;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE P0480E02.2 protein.

```

```

GN P0480E02.2.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
clone: P0480E02.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002913; BAB21196.1; -.
DR Gramene; Q9AWU2; -.
DR InterPro; IPR002057; Isopen_N_synth.
DR InterPro; IPR000270; OPR_PBI.
DR InterPro; IPR002965; P_rich_extensn.
DR Pfam; PF00564; PBI; 1.
DR PRINTS; PR01217; PRICHEXTENS.
DR SMART; SM00666; PBI; 1.
DR PROSITE; PS00185; IPNS_1; 1.
SQ SEQUENCE 731 AA; 77104 MW; 7B00D2CE2798F36B CRC64;

Query Match 2.2%; Score 8; DB 10; Length 731;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 PQQAPQP 112
DB 560 PQQAPQP 567

RESULT 72
O02402 PRELIMINARY; PRT; 738 AA.
AC O02402;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Insoluble protein.
OS Pinctada fucata.
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pterioida;
OC Pterioidea; Pteriidae; Pinctada.
OX NCBI_TaxID=50426;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97320490; PubMed=9177341;
RA Sudo S., Fujikawa T., Nagakura T., Ohkubo T., Sakaguchi K., Tanaka M.,
RA Nakashima K., Takahashi T.;
RT "Structures of mollusc shell framework proteins.";
RL Nature 387:563-564(1997).
DR EMBL; D86074; BAA20466.1; -.
DR InterPro; IPR00104; Antifreeze_1.
DR InterPro; IPR002952; Eggshell.
DR InterPro; IPR000817; Pdon.
DR PRINTS; PR00308; ANTIFREEZE1.
DR PRINTS; PR01228; EGGSHELL.
DR PRINTS; PR00341; PRION.
SQ SEQUENCE 738 AA; 61723 MW; FDF984139BF3BA59 CRC64;

Query Match 2.2%; Score 8; DB 5; Length 738;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGG 241
DB 119 AGGGAGG 126

RESULT 73
O62006 PRELIMINARY; PRT; 795 AA.
ID O62006

```

AC 062006;
DT 01-AUG-1998 (TRENBLrel. 07, Created)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Intermediate filament protein C2.
OS Branchiostoma lanceolatum (Common lancelet) (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
OX NCBI_TaxID=7740;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98267218; PubMed=9602172;
RA Riemer D., Karabinos A., Weber K.;
RT "Analysis of eight cDNAs and six genes for intermediate filament (IF) proteins in the cephalochordate Branchiostoma reveals differences in the IF multi-gene families of lower chordates and the vertebrates.";
RL Gene 211:361-373(1998).
DR EMBL; AJ223578; CAA11445.1; -;
DR InterPro; IPR002952; Eggshell.
DR InterPro; IPR001664; IF.
DR InterPro; IPR002957; Keratin_1.
DR Pfam; PF00038; filament; 5.
DR PRINTS; PRO1228; EGGSHLL.
DR PROSITE; PRO1248; TYPE1KERATIN.
DR PROSITE; PS00226; IF; 1.
SQ SEQUENCE 795 AA; 84532 MW; D00B2665CEACTFFC CRC64;

Query Match 2.2%; Score 8; DB 5; Length 795;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 241
Db 492 AGGAGGG 499
|||||
[1]
PRELIMINARY; PRT; 809 AA.

AC Q43715;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Chloroplast outer envelope membrane protein (OEP75) precursor (OEP75).
DE (OEP75).
GN IAP75.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids 1; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CV, LITTLE MARVEL; TISSUE=Leaf;
RX MEDLINE=95300776; PubMed=7781598;
RA Tranel P.J., Froehlich J., Goyal A., Keegstra K.;
RT "A component of the chloroplastic protein import apparatus is targeted to the outer envelope membrane via a novel pathway.";
RL EMBO J. 14:2436-2446(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95063931; PubMed=7973649;
RA Schnell D.J., Kessler P., Blobel G.;
RT "Isolation of components of the chloroplast protein import machinery.";
RL Science 266:1007-1012(1994).
DR EMBL; X83767; CAA58720.1; -;
DR EMBL; L36858; AAA53275.1; -;
DR InterPro; IPR000184; Bac surfAg_D15.
DR InterPro; IPR005689; IAP75.
DR Pfam; PF01103; Bac_surface_Ag; 1.
DR TIGRFAWS; TIGR00992; 3a0901S03IAP75; 1.
KW Chloroplast; Outer membrane; Transit peptide.

FT TRANSIT 1 131 POTENTIAL.
FT CHAIN 132 809 CHLOROPLASTIC OUTER ENVELOPE MEMBRANE
FT PROTEIN (OEP75).
SQ SEQUENCE 809 AA; 88269 MW; APES1AE75F0617C5 CRC64;

Query Match 2.2%; Score 8; DB 10; Length 809;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 241
Db 90 AGGAGGG 97
|||||
[1]
PRELIMINARY; PRT; 823 AA.

AC Q25343;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE L2759.4.
GN L2759.4.
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Friedlin;
RX MEDLINE=99178987; PubMed=10077609;
RA Myler P.J., Audleman L., deVos T., Hixson G., Kiser P., Lenley C.,
RA Magness C., Rickel E., Sisk E., Sunkin S., Swartzell S., Westlake T.,
RA Bastien P., Fu G., Ivens A., Stuart K.;
RT "Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-coding genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:2902-2906(1999).
DR EMBL; AE001274; AAC24635.1; -;
DR InterPro; IPR006626; PBH1.
DR SMART; SM00710; PBH1; 4.
SQ SEQUENCE 823 AA; 86775 MW; 5631E7DFC1A70435 CRC64;

Query Match 2.2%; Score 8; DB 5; Length 823;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 56 ESTLEENK 63
Db 736 ESTLEENK 743
|||||
[1]
ESTLEENK 743

Search completed: December 17, 2003, 06:35:19
Job time : 41.1811 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 06:20:56 ; Search time 34.7529 Seconds
(without alignments)
1680.762 Million cell updates/sec

Title: US-09-805-550-4

Perfect score: 368

Sequence: 1 MKLTVKTLGTHFPIRVQPN.....CDRNEELAAVYLLHAGEED 368

Scoring table:

Gapop 60.0 ; Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : A.Geneseq 19Jun03.*

```

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	368	100.0	368	21	Maize Rad23 protel
2	56	15.2	117	23	Human ORF391 prot
3	21	5.7	299	21	Arabidopsis thalia
4	21	5.7	348	21	Arabidopsis thalia
5	21	5.7	371	21	Arabidopsis thalia
6	15	4.1	186	21	Arabidopsis thalia
7	15	4.1	214	21	Arabidopsis thalia
8	15	4.1	257	21	Arabidopsis thalia
9	15	4.1	296	21	Arabidopsis thalia

10	15	4.1	307	21	Arabidopsis thalia
11	15	4.1	307	21	Arabidopsis thalia
12	15	4.1	335	21	Arabidopsis thalia
13	15	4.1	335	21	Arabidopsis thalia
14	15	4.1	345	21	Arabidopsis thalia
15	15	4.1	368	21	Arabidopsis thalia
16	15	4.1	378	21	Arabidopsis thalia
17	15	4.1	378	21	Arabidopsis thalia
18	13	3.5	405	21	Arabidopsis thalia
19	11	3.0	136	21	Maize Rad23 protel
20	11	3.0	136	21	Arabidopsis thalia
21	9	2.4	424	22	Drosophila melanog
22	8	2.2	92	22	Novel human diagno
23	8	2.2	10	22	Arabidopsis thalia
24	8	2.2	10	22	Arabidopsis thalia
25	8	2.2	60	23	Human ORF3579 prot
26	8	2.2	174	21	Arabidopsis thalia
27	8	2.2	265	19	FLGA insert stabil
28	8	2.2	325	10	Sequence of the 65
29	8	2.2	331	24	Escherichia tenella ep
30	8	2.2	332	21	Human cancer-relat
31	8	2.2	352	11	Barley NADPH-thior
32	8	2.2	566	22	Antigen GX3271 enc
33	8	2.2	604	21	Drosophila melanog
34	8	2.2	620	22	Mouse retinoid bin
35	8	2.2	712	11	Drosophila melanog
36	8	2.2	738	19	Protein having imm
37	8	2.2	828	23	New DNA sequence i
38	8	2.2	828	24	Human polypeptide
39	8	2.2	1004	22	Human cell adhesio
40	8	2.2	1088	24	C glutamicum prote
41	8	2.2	1329	23	S. cinnamomeus cin
42	8	2.2	1385	22	Human NOV3a protei
43	8	2.2	1417	22	Drosophila melanog
44	8	2.2	2075	22	Drosophila melanog
45	7	1.9	9	22	Drosophila melanog
46	7	1.9	9	22	Human kinase fragm
47	7	1.9	10	22	Human kinase fragm
48	7	1.9	10	22	Human kinase fragm
49	7	1.9	10	22	Human complementar
50	7	1.9	10	22	Human complementar
51	7	1.9	10	22	Human complementar
52	7	1.9	10	22	Human complementar
53	7	1.9	10	22	Arabidopsis thalia
54	7	1.9	10	22	Arabidopsis thalia
55	7	1.9	12	17	Cell adhesion modu
56	7	1.9	14	13	Sequence of synthe
57	7	1.9	14	13	Synthetic random c
58	7	1.9	15	7	Epstein-Barr virus
59	7	1.9	15	24	Human cancer-relat
60	7	1.9	15	24	Human cancer-relat
61	7	1.9	15	24	Human cancer-relat
62	7	1.9	15	24	Human cancer-relat
63	7	1.9	15	24	Human cancer-relat
64	7	1.9	15	24	Human cancer-relat
65	7	1.9	15	24	Human cancer-relat
66	7	1.9	16	13	Sequence of synthe
67	7	1.9	16	13	Synthetic random c
68	7	1.9	17	13	Sequence of synthe
69	7	1.9	17	13	Synthetic random c
70	7	1.9	17	19	Stabilising sequen
71	7	1.9	19	23	Streptococcus pneu
72	7	1.9	20	7	Epstein-Barr virus
73	7	1.9	20	11	Antigenic fragment
74	7	1.9	20	13	Sequence of synthe
75	7	1.9	20	13	Synthetic random c
76	7	1.9	20	13	Polypeptide P-62.
77	7	1.9	24	11	Antigenic fragment
78	7	1.9	25	11	Antigenic fragment
79	7	1.9	35	15	Eukaryotic folding
80	7	1.9	35	15	Human liver peptid
81	7	1.9	55	22	Propionibacterium
82	7	1.9	55	22	Peptide #5611 enco

83 7 1.9 55 22 AB23322 Protein #5321 enco
 84 7 1.9 55 22 AAM58743 Human brain expres
 85 7 1.9 55 22 AAM71250 Human bone marrow
 86 7 1.9 55 22 AAM18952 Peptide #5386 enco
 87 7 1.9 55 22 AAM31531 Peptide #5568 enco
 88 7 1.9 55 23 ABG41049 Human peptide enco
 89 7 1.9 56 23 ABP29010 Streptococcus poly
 90 7 1.9 57 17 AAR88669 Human cytomagalovi
 91 7 1.9 61 19 AAW75013 Human secreted pro
 92 7 1.9 61 23 ABC95475 Human novel secret
 93 7 1.9 61 23 ABP05846 Human ORFX protein
 94 7 1.9 70 21 AAG09626 Arabidopsis thalia
 95 7 1.9 72 23 ABP03541 Human ORFX protein
 96 7 1.9 78 23 ABG60309 Lymphona associate
 97 7 1.9 78 23 ABG60310 Lymphona associate
 98 7 1.9 80 21 AAG57031 Arabidopsis thalia
 99 7 1.9 81 20 AAY32776 HIV SP-2 strain ch
 100 7 1.9 82 20 AAY32775 HIV SP-2 strain ch

ALIGNMENTS

RESULT 1

AA71459
 ID AAY71459 standard; Protein; 368 AA.

AC AAY71459;
 XX

DT 04-OCT-2000 (first entry)
 XX

DE Maize Rad23 protein #2.
 XX

XX Rad23; maize; ATCC No: PTA-531; recombinant expression cassette; wheat;
 KW transgenic plant; soybean; sunflower; canola; modulator.
 XX

OS Zea mays.
 XX

PN WO200031268-A1.
 XX

PD 02-JUN-2000.
 XX

PF 12-OCT-1999; 99WO-US24129.
 XX

PR 23-NOV-1998; 98US-0109728.
 XX

PA (PION-) PIONEER HI-BRED INT INC.
 XX

PI Mahajan PB, Tagliani L;
 XX

DR N-PSDB; AAD01231.
 XX

XX WPI; 2000-400078/34.
 XX

XX Isolated nucleic acid encoding maize RAD23 protein is used to modulate
 PT the levels of polypeptides in plant or in assays for identifying
 PT compounds that bind to and/or increase/decrease enzymatic activity of
 PT catalytically active polypeptides -
 XX

XX Claim 11b; Page 78-79; 82pp; English.
 XX

CC The present sequence is the maize Rad23 protein #2. It is isolated from
 CC a Zea mays cell line, B73 callus tissue regenerated five days after
 CC transfer of the callus from medium containing auxin to a medium devoid
 CC of exogenous auxin. The cDNA is deposited under the ATCC No: PTA-531.
 CC Maize Rad23 DNA sequence operably linked to a promoter, can be used to
 CC construct a recombinant expression cassette. This expression cassette
 CC can be used to generate a dicot or monocot transgenic plant e.g., maize,
 CC soybean, sunflower, sorghum, canola, wheat, etc. It can also be used to
 CC modulate the levels of Rad23 polypeptide expression in a plant or in
 CC assays to identify compounds, that bind to and/or modulate the enzymatic
 CC activity of catalytically active polypeptides.

XX Sequence 368 AA;
 XX

Query Match 100.0%; Score 368; DB 21; Length 368;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLTVKTLKGTHTFEIRVQPNNDTMAVKKNIIEIQKDSYPMGQQLLIENGKVLKDESTLE 60
 DB 1 MKLTVKTLKGTHTFEIRVQPNNDTMAVKKNIIEIQKDSYPMGQQLLIENGKVLKDESTLE 60
 QY 61 ENKVNEDGLVVMLSKGTSGTSSQHSNTPATROAPPLEAPQAPQPPVAPITTSQ 120
 DB 61 ENKVNEDGLVVMLSKGTSGTSSQHSNTPATROAPPLEAPQAPQPPVAPITTSQ 120
 QY 121 PEGLPAPQPNTHDAAASNLLSGRNVDTIINQIMEMGGSGWDKQVQRAIRAAVNNPERAV 180
 DB 121 PEGLPAPQPNTHDAAASNLLSGRNVDTIINQIMEMGGSGWDKQVQRAIRAAVNNPERAV 180
 QY 181 EYLYSGIPVTABIAVPIGGQANTTDRAPTGTAGLSGIPNTAPLDLPFGASNAGGAGG 240
 DB 181 EYLYSGIPVTABIAVPIGGQANTTDRAPTGTAGLSGIPNTAPLDLPFGASNAGGAGG 240
 QY 241 GPLDFLRNNPQAVREVMVHTNPQILQPMVLVSKNPQILRLIENHDEFLLNPEPE 300
 DB 241 GPLDFLRNNPQAVREVMVHTNPQILQPMVLVSKNPQILRLIENHDEFLLNPEPE 300
 QY 301 GGEGDFLOQPEDEMPHAI SVTPPEQEAIGRLSMGFDRARVIEAPLACDRNEELAANYL 360
 DB 301 GGEGDFLOQPEDEMPHAI SVTPPEQEAIGRLSMGFDRARVIEAPLACDRNEELAANYL 360
 QY 361 LEHAGEED 368
 DB 361 LEHAGEED 368

RESULT 2

ABP34618
 ID ABP34618 standard; Protein; 117 AA.

AC ABP34618;
 XX

DT 08-JUL-2002 (first entry)
 XX

DE Human ORF3591 protein, SEQ ID NO: 7182.
 XX

XX Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; haematopoiesis regulation; tissue growth;
 KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
 KW vasotropic; antipsoriatic; antidiabetic; cytostatic; neutropic;
 KW neuroprotective; antilatherosclerotic; anticoagulant; thrombolytic;
 KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide.

OS Homo sapiens.
 XX

XX WO200190366-A2.
 XX

PD 29-NOV-2001.
 XX

PF 24-MAY-2001; 2001WO-US17076.
 XX

PR 24-MAY-2000; 2000US-206690P.
 XX

PA (CURA-) CURAGEN CORP.
 XX

PI Leach MD, Shimkets RA;
 XX

XX WPI; 2002-106200/14.
 DR

DR N-PSDB; ABN78644.
XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and disorders related to organ
PT transplantation -
XX Claim 10; Page 2042; 2508pp; English.
XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
CC ABN79587 represent cDNAs encoding them. The invention also encompasses
CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
CC polynucleotides, the recombinant production of ORFX proteins, antibodies
CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
CC polypeptides, methods of screening for modulators of ORFX expression or
CC activity, and methods of screening individuals for a predisposition to an
CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
CC range of biological activities, such as cytokine, cell proliferation,
CC cell differentiation, immune modulation, haematopoiesis regulation,
CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
CC chemokinetic activity, haemostatic activity, thrombolytic activity,
CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
CC and antifibrotic activity, and may also be involved in the determination
CC of bodily characteristics, fertility and behaviour. ORFX proteins,
CC nucleic acids and antibodies may be used in the treatment of cancers,
CC other proliferative disorders such as epilepsy and benign tumours,
CC neurological disorders such as epilepsy and Alzheimer's disease,
CC cardiovascular diseases, immune system disorders, disorders related to
CC organ transplantation, disorders of tissue growth and regeneration,
CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
CC storage disease, and infectious diseases caused by viral, bacterial,
CC fungal and other pathogens. ORFX nucleic acids may also be used as a
CC source of primers and probes, in the detection of ORFX genomic sequences
CC or transcripts, in the identification and cloning of homologous
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC nucleic acids may additionally be used to produce transgenic animals
CC which may be useful for studying the function and/or activity of ORFX
CC protein, and in drug screening. The ORFX proteins may also be used as
CC immunogens to generate specific antibodies, which are useful in the
CC diagnosis, treatment and monitoring of ORFX-associated diseases.
XX
SQ Sequence 117 AA;
Query Match 15.2%; Score 56; DB 23; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.5e-47;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 253 QAVREMVHTNPQILQPMVLVLSKQNPQILRLIEENHDFLQLNPEFEGSGDFLD 308
Db 29 QAVREMVHTNPQILQPMVLVLSKQNPQILRLIEENHDFLQLNPEFEGSGDFLD 84
RESULT 3
AAG17379
ID AAG17379 standard; Protein; 299 AA.
XX AAG17379;
XX AAG17379;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 18374.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
OS Arabidopsis thaliana.
XX Arabidopsis thaliana.
PN EP1033405-A2.
XX

PD 06-SEP-2000.
XX 25-FEB-1999; 99US-0121825.
XX 25-FEB-1999; 99US-0123180.
XX 05-MAR-1999; 99US-0123548.
XX 09-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 18-MAY-1999; 99US-0134370.
XX 19-MAY-1999; 99US-0134768.
XX 20-MAY-1999; 99US-0134941.
XX 21-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140695.
XX 28-JUN-1999; 99US-0140823.
XX 29-JUN-1999; 99US-0140931.
XX 30-JUN-1999; 99US-0141287.
XX 01-JUL-1999; 99US-0141842.
XX 01-JUL-1999; 99US-0142154.
XX 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.

PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-015753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 5.7%; Score 21; DB 21; Length 299;
Best Local Similarity 100.0%; Pred. No. 4e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 342 VIEAFACDRNEELAANYLLE 362
Db 272 VIEAFACDRNEELAANYLLE 292

RESULT 4

AAG17378
ID AAG17378 standard; Protein; 348 AA.
XX
AC AAG17378;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 18373.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0128264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150586.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.

```
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 5.7%; Score 21; DB 21; Length 348;
Best Local Similarity 100.0%; Pred. No. 4.6e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 342 VIEAFLACDRNEELANYLLE 362
   |||||
Db 321 VIEAFLACDRNEELANYLLE 341

RESULT 5
AAG17377
ID AAG17377 standard; Protein; 371 AA.
XX
AC AAG17377;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 18372.
XX
DE Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
```

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 5.7%; Score 21; DB 21; Length 371;
Best Local Similarity 100.0%; Pred. No. 4.9e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 342 VIEAFLACDRNEELAAANYLLE 362
Db 344 VIEAFLACDRNEELAAANYLLE 364
|||||

RESULT 6

AAG36531
ID AAG36531 standard; Protein; 186 AA.

XX AC AAG36531;

XX 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 44781.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PP 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 08-APR-1999; 99US-0128234.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 28-APR-1999; 99US-0130891.

XX 30-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 30-APR-1999; 99US-0132407.

PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 4.1%; Score 15; DB 21; Length 186;

Best Local Similarity 100.0%; Pred. No. 2.4e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 NNPERAVEVLYSGIP 188
 |||||
 DB 106 NNPERAVEVLYSGIP 120

RESULT 7

AAG36530
 ID AAG36530 standard; Protein; 214 AA.

XX AC AAG36530;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 44780.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX OS Arabidopsis thaliana.

XX FN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 18-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0128845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144894.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 4.1%; Score 15; DB 21; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 174 NNPERAVEYLYSGIP 188
Db 134 NNPERAVEYLYSGIP 148
|||||

RESULT 8
AAG36529
ID AAG36529 standard; Protein; 257 AA.
AC AAG36529;
XX
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 44779.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridization assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145921.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-01511303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 4.1%; Score 15; DB 21; Length 257;
Best Local Similarity 100.0%; Pred. NO. 3.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 NNPERAVEVLYSGIP 188
Db 177 NNPERAVEVLYSGIP 191
|||||

RESULT 9
AAG44343
ID AAG44343 standard; Protein; 296 AA.
XX
AC AAG44343;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55533.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD
XX
XX 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150888.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156565.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 4.1%; Score 15; DB 21; Length 296;
Best Local Similarity 100.0%; Pred. No. 3.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 349 CDRNEELAANYLLEH 363
Db 276 CDRNEELAANYLLEH 290
RESULT 10
AAG19976
ID AAG19976 standard; Protein; 307 AA.
XX AC AAG19976;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 21983.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 21-APR-1999; 99US-0130077.
XX PR 23-APR-1999; 99US-0130510.
XX PR 28-APR-1999; 99US-0130891.
XX PR 30-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 04-MAY-1999; 99US-0132407.
XX PR 05-MAY-1999; 99US-0132484.
XX PR 06-MAY-1999; 99US-0132485.
XX PR 06-MAY-1999; 99US-0132486.
XX PR 07-MAY-1999; 99US-0132487.
XX PR 11-MAY-1999; 99US-0132863.
XX PR 14-MAY-1999; 99US-0134218.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 18-MAY-1999; 99US-0134370.
XX PR 19-MAY-1999; 99US-0134768.
XX PR 20-MAY-1999; 99US-0134941.
XX PR 21-MAY-1999; 99US-0135124.
XX PR 21-MAY-1999; 99US-0135353.
XX PR 24-MAY-1999; 99US-0135629.
XX PR 25-MAY-1999; 99US-0136021.
XX PR 27-MAY-1999; 99US-0136392.
XX PR 28-MAY-1999; 99US-0136782.
XX PR 01-JUN-1999; 99US-0137222.
XX PR 03-JUN-1999; 99US-0137528.
XX PR 04-JUN-1999; 99US-0137502.
XX PR 07-JUN-1999; 99US-0137724.
XX PR 08-JUN-1999; 99US-0138094.
XX PR 10-JUN-1999; 99US-0138540.
XX PR 10-JUN-1999; 99US-0138847.
XX PR 14-JUN-1999; 99US-0139119.
XX PR 16-JUN-1999; 99US-0139452.
XX PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150568.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 4.1%; Score 15; DB 21; Length 307;

Best Local Similarity 100.0%; Pred. No. 3.8e-06; Mismatches 0; Indels 0; Gaps 0;

QY 174 NNPERAVEYLXSGIP 188
Db 106 NNPERAVEYLXSGIP 120

RESULT 11

AA045206	18-OCT-2000	(first entry)	Arabidopsis thaliana protein fragment SEQ ID NO: 56725.	PR 18-JUN-1999;	99US-0139460
ID	AA045206	standard; Protein; 307 AA.		PR 18-JUN-1999;	99US-0139461
XX	XX			PR 18-JUN-1999;	99US-0139462
AC	AC			PR 18-JUN-1999;	99US-0139463
XX	XX			PR 18-JUN-1999;	99US-0139464
XX	XX			PR 18-JUN-1999;	99US-0139750
DT	DT			PR 18-JUN-1999;	99US-0139763
XX	XX			PR 21-JUN-1999;	99US-0139817
DE	DE			PR 22-JUN-1999;	99US-0139889
KW	KW			PR 23-JUN-1999;	99US-0140353
KW	KW			PR 23-JUN-1999;	99US-0140354
KW	KW			PR 24-JUN-1999;	99US-0140695
KW	KW			PR 28-JUN-1999;	99US-0140823
OS	OS			PR 29-JUN-1999;	99US-0140991
XX	XX			PR 30-JUN-1999;	99US-0141287
XX	XX			PR 01-JUL-1999;	99US-0141842
XX	XX			PR 01-JUL-1999;	99US-0142154
XX	XX			PR 02-JUL-1999;	99US-0142055
XX	XX			PR 06-JUL-1999;	99US-0142390
XX	XX			PR 08-JUL-1999;	99US-0142803
XX	XX			PR 09-JUL-1999;	99US-0142920
XX	XX			PR 12-JUL-1999;	99US-0142977
XX	XX			PR 13-JUL-1999;	99US-0143542
XX	XX			PR 14-JUL-1999;	99US-0143624
XX	XX			PR 15-JUL-1999;	99US-0144005
XX	XX			PR 16-JUL-1999;	99US-0144085
XX	XX			PR 16-JUL-1999;	99US-0144086
XX	XX			PR 19-JUL-1999;	99US-0144325
XX	XX			PR 19-JUL-1999;	99US-0144331
XX	XX			PR 19-JUL-1999;	99US-0144332
XX	XX			PR 19-JUL-1999;	99US-0144333
XX	XX			PR 19-JUL-1999;	99US-0144334
XX	XX			PR 20-JUL-1999;	99US-0144352
XX	XX			PR 20-JUL-1999;	99US-0144632
XX	XX			PR 20-JUL-1999;	99US-0144884
XX	XX			PR 21-JUL-1999;	99US-0144814
XX	XX			PR 21-JUL-1999;	99US-0145086
XX	XX			PR 22-JUL-1999;	99US-0145085
XX	XX			PR 22-JUL-1999;	99US-0145087
XX	XX			PR 22-JUL-1999;	99US-0145089
XX	XX			PR 22-JUL-1999;	99US-0145192
XX	XX			PR 23-JUL-1999;	99US-0145145
XX	XX			PR 23-JUL-1999;	99US-0145218
XX	XX			PR 23-JUL-1999;	99US-0145224
XX	XX			PR 26-JUL-1999;	99US-0145276
XX	XX			PR 27-JUL-1999;	99US-0145913
XX	XX			PR 27-JUL-1999;	99US-0145919
XX	XX			PR 27-JUL-1999;	99US-0145919
XX	XX			PR 28-JUL-1999;	99US-0145951
XX	XX			PR 02-AUG-1999;	99US-0146386
XX	XX			PR 02-AUG-1999;	99US-0146388
XX	XX			PR 02-AUG-1999;	99US-0146389
XX	XX			PR 03-AUG-1999;	99US-0147038
XX	XX			PR 04-AUG-1999;	99US-0147204
XX	XX			PR 04-AUG-1999;	99US-0147302
XX	XX			PR 05-AUG-1999;	99US-0147192
XX	XX			PR 05-AUG-1999;	99US-0147260
XX	XX			PR 06-AUG-1999;	99US-0147303
XX	XX			PR 06-AUG-1999;	99US-0147416
XX	XX			PR 09-AUG-1999;	99US-0147493
XX	XX			PR 10-AUG-1999;	99US-0147935
XX	XX			PR 10-AUG-1999;	99US-0148171
XX	XX			PR 11-AUG-1999;	99US-0148319
XX	XX			PR 12-AUG-1999;	99US-0148341
XX	XX			PR 13-AUG-1999;	99US-0148565
XX	XX			PR 13-AUG-1999;	99US-0148684
XX	XX			PR 16-AUG-1999;	99US-0149368
XX	XX			PR 17-AUG-1999;	99US-0149175
XX	XX			PR 18-AUG-1999;	99US-0149426
XX	XX			PR 20-AUG-1999;</	

```
PR 23-AUG-1999; 99US-0149802.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159325.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 4.1%; Score 15; DB 21; Length 307;
Best Local Similarity 100.0%; Pred. No. 3.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 NNPERAVEVLYSGIP 188
Db 106 NNPERAVEVLYSGIP 120
|||||
|||||

RESULT 12
AAG19975
ID AAG19975 standard; Protein; 335 AA.
XX
AC AAG19975;
XX
DT 17-OCT-2000 (first entry)
XX
```

DE Arabidopsis thaliana protein fragment SEQ ID NO: 21982.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX EP1033405-A2.
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 23-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0110891.
PR 28-APR-1999; 99US-011449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899;
PR 23-JUN-1999; 99US-0140353;
PR 23-JUN-1999; 99US-0140354;
PR 24-JUN-1999; 99US-0140695;
PR 28-JUN-1999; 99US-0140823;
PR 29-JUN-1999; 99US-0140991;
PR 30-JUN-1999; 99US-0141287;
PR 01-JUL-1999; 99US-0141842;
PR 01-JUL-1999; 99US-0142154;
PR 02-JUL-1999; 99US-0142055;
PR 06-JUL-1999; 99US-0142390;
PR 08-JUL-1999; 99US-0142803;
PR 09-JUL-1999; 99US-0142920;
PR 12-JUL-1999; 99US-0142977;
PR 13-JUL-1999; 99US-0143542;
PR 14-JUL-1999; 99US-0143624;
PR 15-JUL-1999; 99US-0144005;
PR 16-JUL-1999; 99US-0144085;
PR 16-JUL-1999; 99US-0144086;
PR 19-JUL-1999; 99US-0144325;
PR 19-JUL-1999; 99US-0144331;
PR 19-JUL-1999; 99US-0144332;
PR 19-JUL-1999; 99US-0144333;
PR 19-JUL-1999; 99US-0144334;
PR 19-JUL-1999; 99US-0144335;
PR 20-JUL-1999; 99US-0144352;
PR 20-JUL-1999; 99US-0144632;
PR 20-JUL-1999; 99US-0144884;
PR 21-JUL-1999; 99US-0144814;
PR 21-JUL-1999; 99US-0145086;
PR 21-JUL-1999; 99US-0145088;
PR 22-JUL-1999; 99US-0145085;
PR 22-JUL-1999; 99US-0145087;
PR 22-JUL-1999; 99US-0145089;
PR 22-JUL-1999; 99US-0145192;
PR 23-JUL-1999; 99US-0145145;
PR 23-JUL-1999; 99US-0145218;
PR 23-JUL-1999; 99US-0145224;
PR 26-JUL-1999; 99US-0145276;
PR 27-JUL-1999; 99US-0145913;
PR 27-JUL-1999; 99US-0145918;
PR 27-JUL-1999; 99US-0145919;
PR 28-JUL-1999; 99US-0145921;
PR 02-AUG-1999; 99US-0146386;
PR 02-AUG-1999; 99US-0146388;
PR 02-AUG-1999; 99US-0146389;
PR 03-AUG-1999; 99US-0147038;
PR 04-AUG-1999; 99US-0147284;
PR 04-AUG-1999; 99US-0147302;
PR 05-AUG-1999; 99US-0147192;
PR 05-AUG-1999; 99US-0147260;
PR 06-AUG-1999; 99US-0147303;
PR 06-AUG-1999; 99US-0147416;
PR 09-AUG-1999; 99US-0147493;
PR 09-AUG-1999; 99US-0147935;
PR 10-AUG-1999; 99US-0148171;
PR 11-AUG-1999; 99US-0148319;
PR 12-AUG-1999; 99US-0148341;
PR 13-AUG-1999; 99US-0148565;
PR 13-AUG-1999; 99US-0148684;
PR 16-AUG-1999; 99US-0149368;
PR 17-AUG-1999; 99US-0149175;
PR 18-AUG-1999; 99US-0149426;
PR 20-AUG-1999; 99US-0149722;
PR 20-AUG-1999; 99US-0149723;
PR 20-AUG-1999; 99US-0149929;
PR 23-AUG-1999; 99US-0149902;
PR 23-AUG-1999; 99US-0149930;
PR 25-AUG-1999; 99US-0150566;
PR 26-AUG-1999; 99US-0150884;
PR 27-AUG-1999; 99US-0151065;
PR 27-AUG-1999; 99US-0151066;
PR 27-AUG-1999; 99US-0151080;

PR 30-AUG-1999; 99US-0151303;
PR 31-AUG-1999; 99US-0151438;
PR 01-SEP-1999; 99US-0151930;
PR 07-SEP-1999; 99US-0152363;
PR 10-SEP-1999; 99US-0153070;
PR 13-SEP-1999; 99US-0153758;
PR 15-SEP-1999; 99US-0154018;
PR 16-SEP-1999; 99US-0154039;
PR 20-SEP-1999; 99US-0154779;
PR 22-SEP-1999; 99US-0155139;
PR 23-SEP-1999; 99US-0155486;
PR 24-SEP-1999; 99US-0155659;
PR 28-SEP-1999; 99US-0156458;
PR 29-SEP-1999; 99US-0156596;
PR 04-OCT-1999; 99US-0157117;
PR 06-OCT-1999; 99US-0157753;
PR 07-OCT-1999; 99US-0158029;
PR 08-OCT-1999; 99US-0158232;
PR 12-OCT-1999; 99US-0158369;
PR 13-OCT-1999; 99US-0159293;
PR 13-OCT-1999; 99US-0159294;
PR 14-OCT-1999; 99US-0159295;
PR 14-OCT-1999; 99US-0159329;
PR 14-OCT-1999; 99US-0159330;
PR 14-OCT-1999; 99US-0159331;
PR 14-OCT-1999; 99US-0159637;
PR 14-OCT-1999; 99US-0159638;
PR 18-OCT-1999; 99US-0159584;
PR 21-OCT-1999; 99US-0160741;
PR 21-OCT-1999; 99US-0160767;
PR 21-OCT-1999; 99US-0160768;
PR 21-OCT-1999; 99US-0160770;
PR 21-OCT-1999; 99US-0160814;
PR 21-OCT-1999; 99US-0160815;
PR 22-OCT-1999; 99US-0160980;
PR 22-OCT-1999; 99US-0160981;
PR 22-OCT-1999; 99US-0160989;
PR 25-OCT-1999; 99US-0161404;
PR 25-OCT-1999; 99US-0161405;
PR 26-OCT-1999; 99US-0161406;
PR 26-OCT-1999; 99US-0161359;
PR 26-OCT-1999; 99US-0161360;
PR 26-OCT-1999; 99US-0161361;
PR 28-OCT-1999; 99US-0161920;
PR 28-OCT-1999; 99US-0161992;
PR 28-OCT-1999; 99US-0161993;
PR 29-OCT-1999; 99US-0162142;

Query Match 4.1%; Score 15; DB 21; Length 335;
Best Local Similarity 100.0%; Pred. No. 4.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 174 NNPERAVEVLYSGIP 188
Db 134 NNPERAVEVLYSGIP 148
|||||

RESULT 13

AAG45205
ID AAG45205 standard; Protein; 335 AA.

XX AAG45205;

AC AAG45205;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 56724.

XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.

XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132048.
PR 05-MAY-1999; 99US-0132407.
PR 06-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.

```
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159299.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 4.1%; Score 15; DB 21; Length 335;
Best Local Similarity 100.0%; Pred. No. 4.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 174 NNPERAVEYLYSGIP 188
Db 134 NNPERAVEYLYSGIP 148

RESULT 14
AAG44342
ID AAG44342 standard; Protein; 345 AA.
AC AAG44342;
XX
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 55532.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
```

```
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
```

PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.

PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 4.1%; Score 15; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 349 CDRNEELAANYLLEH 363
DB 325 CDRNEELAANYLLEH 339
|||||

RESULT 15
AAG44341

ID AAG44341 standard; Protein; 368 AA.

XX AAG44341;

XX 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 55531.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137503.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.

```
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161520.
PR 28-OCT-1999; 99US-0161592.
PR 28-OCT-1999; 99US-0161593.
PR 29-OCT-1999; 99US-0162142.

Query Match      4.1%; Score 15; DB 21; Length 368;
Best Local Similarity 100.0%; Pred. No. 4.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 349 CDRNEELAAAYLLEH 363
Db 348 CDRNEELAAAYLLEH 362
|||||
|||||

RESULT 16
AAG19974
ID AAG19974 standard; Protein; 378 AA.
XX AC AAG19974;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 21981.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 23-MAR-1999; 99US-0126264.
XX PR 23-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
```

PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155133.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158233.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.

PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 4.1%; Score 15; DB 21; Length 378;
Best Local Similarity 100.0%; Pred. No. 4.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 174 NNPERAVEYLYSGIP 188
|||||
Db 177 NNPERAVEYLYSGIP 191
|||||

RESULT 17

AAG45204
ID AAG45204 standard; Protein; 378 AA.

XX AC AAG45204;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 56723.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX FN EPI033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 21-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150586.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158212.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.

PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 4.1%; Score 15; DB 21; Length 378;

Best Local Similarity 100.0%; Pred. No. 4.6e-06; Indels 0; Gaps 0;
 Matches 15; Conservative 0; Mismatches 0;

Qy 174 NNPERAVEYLYSGIP 188
 |||||
 Db 177 NNPERAVEYLYSGIP 191

RESULT 18

AAV71458
 ID AAY71458 standard; Protein; 405 AA.

AC AAV71458;
 XX
 DT 04-OCT-2000 (first entry)

DE Maize Rad23 protein #1.

XX Rad23; maize; ATCC No: PTA-530; recombinant expression cassette; wheat;
 KW transgenic plant; soybean; sunflower, sorghum; canola; modulator.

XX Zea mays.

OS WO200031268-A1.

XX 02-JUN-2000.

XX 12-OCT-1999; 99WO-US24129.

XX 23-NOV-1998; 98US-0109728.

XX (PION-) PIONEER HI-BRED INT INC.

XX Mahajan PB, Tagliani L;

XX WPI; 2000-400078/34.

XX N-PSDB; RAD01230.

XX Isolated nucleic acid encoding maize RAD23 protein is used to modulate
 PT the levels of polypeptides in plant or in assays for identifying
 PT compounds that bind to and/or increase/decrease enzymatic activity of
 PT catalytically active polypeptides -

XX Claim 11b; Page 75-76; 82pp; English.

XX The present sequence is the maize Rad23 protein #1. It is isolated from
 CC V5 root tissue of a Zea mays cell line B73, infested with corn root
 CC worm. The cDNA sequence is deposited under the ATCC No: PTA-530. Maize
 CC Rad23 DNA sequence operably linked to a promoter can be used to construct
 CC a recombinant expression cassette. This expression cassette can be used
 CC to generate a dicot or monocot transgenic plant e.g., maize, soybean,
 CC sunflower, sorghum, canola, wheat, etc.. It can also be used to modulate
 CC the levels of Rad23 polypeptide expression in a plant or in assays to
 CC identify compounds, that bind to and/or modulate the enzymatic activity
 CC of catalytically active polypeptides.

XX Sequence 405 AA;

Query Match 3.5%; Score 13; DB 21; Length 405;
 Best Local Similarity 100.0%; Pred. No. 0.00048;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 167 RALRAAYNNPERA 179
 |||||
 Db 182 RALRAAYNNPERA 194

RESULT 19

AAG33942
 ID AAG33942 standard; Protein; 136 AA.

XX AAG33942;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 41222.

XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 21-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 28-APR-1999; 99US-0130891.

XX 30-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 04-MAY-1999; 99US-0132407.

XX 05-MAY-1999; 99US-0132484.

XX 06-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 07-MAY-1999; 99US-0132487.

XX 11-MAY-1999; 99US-0132863.

XX 14-MAY-1999; 99US-0134256.

XX 14-MAY-1999; 99US-0134218.

XX 14-MAY-1999; 99US-0134219.

XX 14-MAY-1999; 99US-0134321.

XX 18-MAY-1999; 99US-0134370.

XX 19-MAY-1999; 99US-0134768.

XX 20-MAY-1999; 99US-0134941.

XX 21-MAY-1999; 99US-0135124.

XX 24-MAY-1999; 99US-0135353.

XX 25-MAY-1999; 99US-0135629.

XX 27-MAY-1999; 99US-0136021.

XX 28-MAY-1999; 99US-0136392.

XX 01-JUN-1999; 99US-0136782.

XX 03-JUN-1999; 99US-0137222.

XX 04-JUN-1999; 99US-0137528.

XX 07-JUN-1999; 99US-0137724.

XX 08-JUN-1999; 99US-0138094.

XX 10-JUN-1999; 99US-0138540.

XX 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140595.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140891.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143362.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 08-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 3.0%; Score 11; DB 21; Length 136;
Best Local Similarity 100.0%; Pred. NO. 0.017;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 VKTLKGTHFEI 15
|||||
DB 5 VKTLKGTHFEI 15

XX PN WO200142279-A2.
XX PD 14-JUN-2001.
XX PF 13-DEC-2000; 2000WO-GB04781.
XX PR 13-DEC-1999; 99GB-0029469.
XX PA (PROT-) PROTEOM LTD.
XX PI Roberts GW, Heal JR;
XX DR WPI; 2001-381629/40.
XX PT A set of peptide ligands for agricultural research and development of
PT therapeutic agents comprise specific complementary peptides to proteins
PT encoded by genes of plant genomes -
XX PS Example 4; Page 31; 201pp; English.
XX CC The present invention relates to a set of peptide ligands consisting of
CC specific complementary peptides to proteins encoded by genes of plant
CC genomes. The present sequence is one such peptide from Arabidopsis
CC thaliana. The peptides of the present invention are useful in an assay to
CC identify a peptide, especially a peptide pesticide or herbicide. The
CC peptides are also useful for tools for agricultural research and
CC development.
XX SQ Sequence 10 AA;
Query Match 2.2%; Score 8; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 234 AGGGAGGG 241
DB 1 AGGGAGGG 8
RESULT 23
AAG83368
ID AAG83368 standard; Peptide; 10 AA.
XX AC AAG83368;
XX DT 11-SEP-2001 (first entry)
XX DE Arabidopsis thaliana peptide ligand #8.
XX KW Plant; peptide pesticide; peptide herbicide; agricultural research.
XX OS Arabidopsis thaliana.
XX PN WO200142279-A2.
XX PD 14-JUN-2001.
XX PF 13-DEC-2000; 2000WO-GB04781.
XX PR 13-DEC-1999; 99GB-0029469.
XX PA (PROT-) PROTEOM LTD.
XX PI Roberts GW, Heal JR;
XX DR WPI; 2001-381629/40.
XX CC A set of peptide ligands for agricultural research and development of
PT therapeutic agents comprise specific complementary peptides to proteins
PT encoded by genes of plant genomes -
XX PS Example 4; Page 32; 201pp; English.

XX CC The present invention relates to a set of peptide ligands consisting of
CC specific complementary peptides to proteins encoded by genes of plant
CC genomes. The present sequence is one such peptide from Arabidopsis
CC thaliana. The peptides of the present invention are useful in an assay to
CC identify a peptide, especially a peptide pesticide or herbicide. The
CC peptides are also useful for tools for agricultural research and
CC development.
XX SQ Sequence 10 AA;
Query Match 2.2%; Score 8; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 234 AGGGAGGG 241
DB 1 AGGGAGGG 8
RESULT 24
ABP34606
ID ABP34606 standard; Protein; 60 AA.
XX AC ABP34606;
XX DT 08-JUL-2002 (first entry)
XX DE Human ORF3579 protein, SEQ ID NO:7158.
XX KW Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
KW disease monitoring; cytokine; cell proliferation; cell differentiation;
KW immune modulation; haematopoiesis regulation; tissue growth;
KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
KW behaviour; cancer; proliferative disorder; neurological disorder;
KW cardiovascular disease; immune system disorder; organ transplantation;
KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
KW vasotropic; antipeptidic; antidiabetic; cytostatic; neurotropic;
KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
KW cardiatic; hypotensive; antithyroid; antiinflammatory; immunomodulator;
KW dermatological; analgesic; virucide; antibacterial; fungicide.
XX OS Homo sapiens.
XX PN WO200190366-A2.
XX PD 29-NOV-2001.
XX PF 24-MAY-2001; 2001WO-US17076.
XX PR 24-MAY-2000; 2000US-206690P.
XX PA (CURA-) CURAGEN CORP.
XX PI Leach WD, Shinkets RA;
XX DR WPI; 2002-106200/14.
XX DR N-PSDB; ABN78632.
XX PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and disorders related to organ
transplantation -
XX PS Claim 10; Page 2036; 2508pp; English.
XX CC Sequences ABP1028-ABP35561 represent 4534 novel human proteins
CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
CC ABN79587 represent cDNAs encoding them. The invention also encompasses
CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
referred to as ORFX) proteins, polynucleotides at least 85% identical to

CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, haematopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
 CC and antiinfective activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 CC storage disease, and infectious diseases caused by viral, bacterial,
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.

XX Sequence 60 AA;

Query Match 2.2%; Score 8; DB 23; Length 60;

Best Local Similarity 100.0%; Pred. No. 7.3;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 TPATRQAP 100

|||||||

Db 21 TPATRQAP 28

RESULT 25

AAG39747

ID AAG39747 standard; Protein; 174 AA.

XX AC AAG39747;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 49227.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 09-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0125788.

XX PR 25-MAR-1999; 99US-0126264.

XX PR 29-MAR-1999; 99US-0126785.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 08-APR-1999; 99US-0128714.

XX PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142154.
 PR 06-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.2%; Score 8; DB 21; Length 174;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
|||||
DB 85 AGGGAGGG 92

RESULT 26
AAW79128

ID AAW79128 standard; Protein; 265 AA.

XX AAW79128;

XX DT 18-NOV-1998 (first entry)

DE FLGA insert stabilising polypeptide.

KW Fusion protein; stabilising polypeptide; proteolytic degradation;
KW resistance; half-life; autoimmune disease; inflammation; nitro drug;
KW IkappaB regulator protein; inflammatory bowel disease; in vivo imaging;
KW nitroreductase protein; enzyme therapy; prodrug therapy; protease;
KW cancer; pathological condition.

XX OS Epstein-barr virus.

XX WO9822577-A1.

XX PD 28-MAY-1998.

XX PF 17-NOV-1997; 97WO-IB01508.

XX PR 25-JUN-1997; 97US-0048945.

XX PR 15-NOV-1996; 96US-0030986.

XX PA (MASU/) MASUCCI M G.

XX PI Masucci MG;

XX DR WPI; 1998-312463/27.

XX DR N-PSDB; AAV55630.

PT New fusion proteins resistant to proteolytic degradation -
PT comprising a core protein with a stabilising polypeptide comprising
PT a peptide sequence containing glycine repeats
XX Disclosure; Fig 3; 120pp; English.
PS

XX This represents a stabilising polypeptide and is the FLGA insert of the
 CC invention. The invention provides a method for increasing the resistance
 CC of a core protein to proteolytic degradation that comprises linking or
 CC inserting onto or into the core protein a stabilising polypeptide of
 CC formula (Glya)X(Glyb)Y(Glyc)Z_n where Glya, Glyb, Glyc are 1-6
 CC sequential Gly residues and X, Y, Z are Ala, Ser, val, Ile, Leu, Met,
 CC Phe, Pro or Thr and n can be anything between 1-66. X, Y and Z need not
 CC be identical from n repeat to n repeat. Alternatively a nucleic acid
 CC encoding the stabilising polypeptide can be linked onto or inserted into
 CC a nucleic acid encoding a core protein. The fusion proteins of the
 CC invention are more resistant to degradation by proteases and, thus, have
 CC a longer half-life than the unfused core protein. The products can be
 CC used for treating autoimmune diseases, cancer and inflammation. In
 CC particular, the core protein may be an IkappaB regulator protein for the
 CC treatment of inflammatory bowel disease, or a nitroreductase protein
 CC which can activate nitro drugs in enzyme/prodrug therapy to treat cancer
 CC or other pathological conditions. The fusion proteins can also be used in
 CC diagnostic methods such as in vivo imaging.

SQ Sequence 265 AA;

Query Match 2.2%; Score 8; DB 19; Length 265;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGGG 241
 Db 78 AGGGAGGG 85
 |||||

RESULT 27

AAAP93704
 ID AAP93704 standard; Protein; 325 AA.

XX AC AAP93704;

XX DT 25-MAR-2003 (updated)

XX DT 15-MAR-1992 (first entry)

XX DE Sequence of the 65kd surface antigen recognised by monoclonal
 antibodies 7D1, 7D4 and 20C6.

XX KW Vaccine; coccidiosis; poultry; bivalent vaccine.

XX OS Eimeria.

XX PN EP344808-A.

XX PD 06-DEC-1989.

XX PF 02-JUN-1980; 80EP-0110056.

XX PR 03-JUN-1988; 88US-0202721.

XX PA (HOFF) HOFFMANN-LA ROCHE AG.

XX PI Altenburger W, Binger MH, Chizzonite RA, Kramer RA;

XX PI Lomedico PT, McAndrew SJ;

XX DR WPI; 1989-358220/49.

XX DR N-PSDB; AAN92577.

XX PT DNA sequences encoding Eimeria surface antigens - used in
 PT pox:virus vectors as a vaccine to protect chicks against
 PT coccidiosis

XX PS Claim 3; Fig 17; 78pp; English.

XX CC The inventors claim a new protein which comprises one or more
 CC immunoreactive and/or antigenic determinants of an Eimeria surface
 CC antigen of mol. wt. 28,37,120 or more than 200 kD which specifically
 CC binds to one or more monoclonal antibody (MaB) from ATCC HB 9707-9712

CC (see AAP93703-6). Also new are DNA encoding the protein (see
 CC AAN92576-9), and a vaccine comprising one or more proteins. Vaccine
 CC utility can be enhanced by inserting additional genes into the carrier
 CC virus (see AAP91652).
 CC (Updated on 25-MAR-2003 to correct PF field.)

SQ Sequence 325 AA;

Query Match 2.2%; Score 8; DB 10; Length 325;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 233 NAGGGAGG 240
 |||||

Db 225 NAGGGAGG 232
 |||||

RESULT 28

AAW31583
 ID AAW31583 standard; Protein; 325 AA.

XX AC AAW31583;

XX DT 25-MAR-2003 (updated)

XX DT 21-MAY-1998 (first entry)

XX DE Eimeria tenella sporozoite, schizont, merozoite antigen.

XX KW Coccidiosis; vaccine; poultry; protozoan; parasite; antigen;

XX KW sporozoite; schizont; merozoite.

XX OS Eimeria tenella.

XX PH Key Location/Qualifiers

FT Peptide 1..20

FT /note= "the first approx. 20 amino acids of the
 FT sequence may result from an artifact during
 FT cDNA synthesis which caused a looping back
 FT of a more internal sequence"

XX PN US5661015-A.

XX PD 26-AUG-1997.

XX PF 20-DEC-1991; 91US-0812349.

XX PR 20-DEC-1991; 91US-0812349.

XX PR 03-JUN-1988; 88US-0202721.

XX PA (HOFF) HOFFMANN LA ROCHE INC.

XX PI Altenburger W, Binger M, Chizzonite RA, Kramer RA;

XX PI Lomedico PT, McAndrew SJ;

XX DR WPI; 1997-434379/40.

XX DR N-PSDB; AAT93594.

XX PT New DNA from Eimeria tenella and related immunogenic polypeptides -
 PT useful in vaccines to protect poultry against coccidiosis

XX PS Example 6.5; Fig 17; 72pp; English.

XX CC This 65 kDa protein is recognised by monoclonal antibodies 7D1,
 CC 20C6 and 7D4 (ATCC HB 9707). These antibodies also specifically
 CC react with an Eimeria tenella 120 kDa surface antigen that is
 CC present in the sporozoite, schizont and merozoite developmental
 CC stages. The amino acid sequence was deduced from a cDNA clone
 CC (see AAT93594) obtained from a cDNA library by immunological
 CC screening with monoclonal antibodies raised against Eimeria
 CC antigens. The invention provides DNA sequences (see AAT93593-98)
 CC coding for Eimeria surface antigens (see AAW31582-84 and AAW33621-26),
 CC recombinant vectors containing such DNA sequences, transformed
 CC microorganisms containing such vectors, and methods for producing

CC the antigens using the transformed microorganisms. Methods are
 CC also provided for protecting poultry against coccidiosis using the
 CC Eimeria surface antigens. The surface antigens are administered
 CC either as purified proteins or in the form of DNA encoding the
 CC proteins in a viral vector such as a vaccinia virus. The vaccines
 CC may produce antibodies that are cross-reactive with other Eimeria
 CC species.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX
 XX

SQ Sequence 325 AA;

Query Match 2.2%; Score 8; DB 18; Length 325;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 NAGGGAGG 240

DB 225 NAGGGAGG 232

RESULT 29

ABR01845

ID ABR01845 standard; Protein; 331 AA.

XX
 AC ABR01845;

DT 19-MAY-2003 (first entry)

DE Human cancer-related protein, 161P2B7A v.3.

XX Human; cytostatic; vaccine; cancer; immune response.

XX Homo sapiens.

PN WO200283921-A2.

XX 24-OCT-2002.

PF 10-APR-2002; 2002WO-US11654.

PR 10-APR-2001; 2001US-282739P.

PR 10-APR-2001; 2001US-283112P.

PR 25-APR-2001; 2001US-286630P.

XX (AGEN-) AGENSYS INC.

XX Jakobovits A, Challita-Bid PM, Faris M, Ge W, Hubert RS;
 PI Morrison K, Morrison RK, Raitano AB;

XX WPI; 2003-075555/07.

DR N-PSDB; ABZ78155.

XX New composition comprising a substance that modulates the structure of
 PT proteins and polynucleotides, useful for therapeutic, prognostic and
 PT diagnostic reagents for eliciting cellular or humoral immune response
 in cancer patients -
 PT
 XX

PS Example 5; Fig 12j; 1021pp; English.

XX The present invention relates to novel human cancer-related genes and
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer.
 XX

SQ Sequence 331 AA;

Query Match 2.2%; Score 8; DB 24; Length 331;

Best Local Similarity 100.0%; Pred. No. 37;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241

DB 83 AGGGAGGG 90

RESULT 30

AAB29250

ID AAB29250 standard; Protein; 332 AA.

XX
 AC AAB29250;

DT 07-FEB-2001 (first entry)

DE Barley NADPH-thioredoxin reductase.

XX Barley; NADPH-thioredoxin reductase; disulfide bridge reduction; NTR;
 KW development; thioredoxin h; hair care product; venom neutralisation;
 KW food technology; food allergy.

XX Hordeum vulgare.

OS WO200058352-A2.

XX 05-OCT-2000.

PF 31-MAR-2000; 2000WO-US08566.

XX 31-MAR-1999; 99US-0127198.

PR 06-DEC-1999; 99US-0169162.

PR 21-JAN-2000; 2000US-0177739.

XX 21-JAN-2000; 2000US-0177740.

XX (REGC) UNIV CALIFORNIA.

XX Cho M, Del Val G, Caillaud M, Lemaux PG, Buchanan BB;
 PI WPI; 2000-679291/66.

DR N-PSDB; AAC62460.

XX Recombinant or isolated nucleic acid, useful for producing transgenic
 PT plants with altered redox properties, encode thioredoxin h or
 PT thioredoxin reductase -
 PT
 XX

PS Claim 49; Fig 4; 125pp; English.

XX The present invention relates to the isolation and use of the barley
 CC thioredoxin h and NADPH-thioredoxin reductase coding and protein
 CC sequences. Thioredoxin is thought to be involved in plant development via
 CC its function in the reduction of disulfide bridges. Thioredoxin can be
 CC used in hair care products and in the neutralisation of some venoms and
 CC toxins, and is also useful in the reduction of some food, for example it
 CC can be used to reduce the allergenicity of foods and the digestibility
 CC of some proteins. It can also be used to enhance the baking qualities of
 CC cereal flour.
 XX

SQ Sequence 332 AA;

Query Match 2.2%; Score 8; DB 21; Length 332;

Best Local Similarity 100.0%; Pred. No. 37;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 GGAGGGGGL 243

DB 221 GGAGGGGGL 228

RESULT 31

AAR05217

ID AAR05217 standard; protein; 352 AA.

XX

AC AAR05217;
 XX 25-MAR-2003 (updated)
 DT 02-AUG-1990 (first entry)
 XX
 DE Antigen GX3271 encoded by cDNA of Eimeria tenella.
 XX
 KW Eimeria tenella; antigen GX3271; avian coccidiosis.
 XX
 OS Eimeria tenella.
 XX
 PN WO9000403-A.
 XX
 XX 25-JAN-1990.
 PD
 XX
 PF 05-JUL-1989; 89WO-US02918.
 XX
 XX 05-JUL-1988; 88US-0215162.
 PR
 XX (GEMX) GENEX CORP.
 PA
 XX Anderson DM, McCandless RJ, Strausberg SL, Strausberg RL;
 PI
 XX WPI; 1990-051586/07.
 DR
 XX N-PSDB; AAQ03319.
 DR
 XX Cloned gene or fragment encoding antigenic protein -
 PT which binds with antibodies against avian coccidia, and
 PT transformed cells used in vaccine
 PT
 XX Claim 5; Page 93; Fig 3; 134pp; English.
 PS
 XX An E. tenella cDNA library was screened with chicken immune serum. Two
 CC phage plaques were identified that produce coccidial antigens reactive
 CC with this chicken immune serum. The coccidial DNA from these plaques was
 CC transferred to bacteriophage M13 for DNA sequence analysis. Their DNA
 CC sequences do not show homology with each other, nor with the GX5401
 CC (AA93324) coding sequence. Plasmid expression vectors were constructed
 CC for the production of them in E. coli. They were designated PGX3271 and
 CC PGX3273, and encode beta-galactosidase coccidial Ag. fusion proteins.
 CC Also new are host cells transformed with the vectors. The transformed
 CC cells are used in vaccine to immunise birds against avian coccidiosis.
 CC By labelling the peptides, they can be used as a type-specific probe.
 CC The AP may also be used in an assay to detect Ab against the coccidia.
 CC The Abs are used to identify transformed cells contg. the DNA.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 352 AA;
 Query Match 2.2%; Score 8; DB 11; Length 352;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 233 NAGGGAGG 240
 Db 252 NAGGGAGG 259
 |||||
 RESULT 32
 ABB65875
 ID ABB65875 standard; Protein; 566 AA.
 XX
 AC ABB65875;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 24417.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX

PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI
 XX WPI; 2001-656860/75.
 DR
 XX N-PSDB; ABL09978.
 DR
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PT
 XX Disclosure; SEQ ID NO 24417; 21pp + Sequence Listing; English.
 PS
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABBS7737-ABBS72072).
 CC
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 566 AA;
 SQ
 Query Match 2.2%; Score 8; DB 22; Length 566;
 Best Local Similarity 100.0%; Pred. No. 62;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 234 AGGGAGGG 241
 Db 95 AGGGAGGG 102
 |||||
 RESULT 33
 AAY57087
 ID AAY57087 standard; protein; 604 AA.
 XX
 AC AAY57087;
 XX
 DT 28-FEB-2000 (first entry)
 XX
 DE Mouse retinoid binding protein.
 XX
 XX Neglected target tissue antigen; NNTA; autoimmunity; autoimmune response;
 KW immunotherapeutic agent; insulin dependent diabetes mellitus;
 KW multiple sclerosis; autoimmune thyroiditis; rheumatoid arthritis;
 KW uveoretinitis; inflammatory response.
 XX
 OS Mus sp.
 XX
 XX WO9956763-A1.
 PN
 XX 11-NOV-1999.
 PD
 XX 07-MAY-1999; 99WO-US10250.
 PF
 XX 07-MAY-1998; 98US-0084636.
 PR
 XX (REGC) UNIV CALIFORNIA.
 PA
 XX Kaufman DL, Tian J, Olcott A;
 PI

XX WIPI; 2000-052905/04.
 XX Administration of neglected target tissue antigens to modulate immune
 XX responses -
 PT
 XX Disclosure; Page 29; 79pp; English.
 XX
 CC Amino acid sequences AAY57063-Y57091 are examples of neglected target
 CC tissue antigens NTAs. NTAs are antigens (whole antigens or fragments)
 CC not involved in autoimmunity. These peptides and proteins are used in
 CC the method of the invention which involves administering an NTA as an
 CC antigen based immunotherapeutic agent, to a host afflicted with an
 CC autoimmune response associated with an autoimmune disease. The
 CC immunotherapeutic agent is used to treat autoimmune diseases such as
 CC insulin dependent diabetes mellitus, multiple sclerosis, autoimmune
 CC thyroiditis, uveoretinitis, rheumatoid arthritis or abnormal
 CC inflammatory immune responses. The NTA induces regulatory tolerance by
 CC elicitation of regulatory T cells among T cells recognizing the NTA but
 CC not participating in the immune response. The NTA are capable of
 CC recognition by substantial populations of uncommitted T cells which can
 CC be primed, or biased, towards regulatory responses to provide effective
 CC treatment. The NTA are effective in regulating undesirable immune
 CC responses even when target determinants used as agents promoting
 CC tolerance. NTAs have failed to induce an effective regulatory T cell
 CC response. NTAs as agents promoting tolerance are anticipated to be safer
 CC than use of target determinants.
 XX
 XX Sequence 604 AA;
 SQ
 Query Match 2.2%; Score 8; DB 21; Length 604;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 102 LEAPQAP 109
 DB 90 LEAPQAP 97
 |||||
 |||||
 RESULT 34
 ABB69938
 ID ABB69938 standard; Protein; 620 AA.
 XX
 AC ABB69938;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 36606.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WIPI; 2001-656860/75.
 DR N-PSDB; ABL14041.
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

XX Disclosure; SEQ ID NO 36606; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 620 AA;
 SQ
 Query Match 2.2%; Score 8; DB 22; Length 620;
 Best Local Similarity 100.0%; Pred. No. 68;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 234 AGGGAGGG 241
 DB 26 AGGGAGGG 33
 |||||
 |||||
 RESULT 35
 AAR07094
 ID AAR07094 standard; protein; 712 AA.
 XX
 AC AAR07094;
 XX
 DT 25-MAR-2003 (updated)
 DT 24-JAN-1991 (first entry)
 XX
 DE Protein having immunological properties of E.tenella.
 XX
 KW Coccidiosis; vaccine; monoclonal antibodies.
 XX
 OS Eimeria tenella.
 XX
 PN EP390267-A.
 XX
 PD 03-OCT-1990.
 XX
 PF 22-MAR-1990; 90EP-0200686.
 XX
 PR 22-MAR-1990; 90EP-0200686.
 XX
 PA (ALKU) AKZO NV.
 PA (ALKU) AKZO NOBEL NV.
 XX
 PI Dijkema R, Vermeulen A, Clarke LE, Tomly FM;
 XX
 DR WIPI; 1990-299216/40.
 DR N-PSDB; AAQ06033.
 XX
 PT Coccidiosis vaccine protein - obtained by immune adsorption of
 PT eimeria tenella extract to substrate containing monoclonal
 PT antibody and releasing adsorption fraction
 XX
 PS Disclosure; fig 6; 28pp; English.
 XX
 CC This E.tenella protein can be used in bird vaccines for protection
 CC against Coccidiosis. Constituent polypeptides can also be used,
 CC or antibodies raised against them. See also AAR07092-93.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 XX Sequence 712 AA;
 SQ
 Query Match 2.2%; Score 8; DB 11; Length 712;
 Best Local Similarity 100.0%; Pred. No. 77;

```
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 NAGGAGG 240
   |||||
Db 612 NAGGAGG 619

RESULT 36
AAW56163
ID AAW56163 standard; Protein; 738 AA.
XX
AC AAW56163;
XX
DT 28-JUL-1998 (first entry)
XX
DE New DNA sequence isolated from Pinctada fucata.
XX
KW Pinctada fucata; protein; cosmetic.
XX
OS Pinctada fucata.
XX
PN JP10080285-A.
XX
PD 31-MAR-1998.
XX
PF 28-MAY-1997; 97JP-0138461.
XX
PR 15-JUL-1996; 96JP-0184459.
XX
PA (MIKI-) MIKIMOTO SEIYAKU KK.
XX
XX WPI; 1998-254410/23.
DR N-PSDB; AAW22683.
XX
PT New cDNA and e.g. vector, host cell and polypeptide - used to
PT produce polypeptide in high yields, which is used in cosmetics
XX
PS Claim 9; Pages 9-11; 15pp; Japanese.
XX
CC The present sequence represents protein encoded by a new DNA sequence
CC isolated from Pinctada fucata. The protein be used as an ingredient
CC in cosmetics.
XX
SQ Sequence 738 AA;

Query Match 2.2%; Score 8; DB 19; Length 738;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
   |||||
Db 119 AGGGAGGG 126

RESULT 37
ABP69286
ID ABP69286 standard; Protein; 828 AA.
XX
AC ABP69286;
XX
DT 20-JAN-2003 (first entry)
XX
DE Human polypeptide SEQ ID NO 1333.
XX
KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW haemostatic; vulnary; fungicide; antibacterial; virucide; protozoacide;
KW antiarthritic.
XX
```

```
OS Homo sapiens.
XX
PN WO200270539-A2.
XX
PD 12-SEP-2002.
XX
PF 05-MAR-2002; 2002WO-US05095.
XX
PR 05-MAR-2001; 2001US-0799451.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren P;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX
DR WPI; 2002-759812/82.
DR N-PSDB; ABZ11503.
XX
XX New polynucleotides comprising sequences assembled from expressed
XX sequence tags (ESTs), useful for treating cell-proliferative,
XX neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or
XX platelet or coagulation disorders -
XX
PS Claim 9; SEQ ID NO 1333; 1012pp + Sequence Listing; English.
XX
XX The invention relates to an isolated polynucleotide (I) comprising a
XX nucleotide sequence selected from any of 948 sequences
XX (ABZ11119-ABZ12066) or their mature protein coding portion, active domain
XX coding protein or complementary sequences. The polynucleotides are useful
XX for identifying expressed genes or for physical mapping of human genome.
XX The encoded polypeptides (ABP68902-ABP69849) are useful as molecular
XX weight markers, as a food supplement, for generating antibodies, in
XX medical imaging, screening and diagnostic assays and for treating
XX cell-proliferative disorders (cancer), neurodegenerative diseases
XX (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple
XX sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid
XX disorders, platelet or coagulation disorders, wound, burns, incision,
XX ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,
XX parasitic), arthritis, etc.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 828 AA;

Query Match 2.2%; Score 8; DB 23; Length 828;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
   |||||
Db 706 AGGGAGGG 713

RESULT 38
ABP58225
ID ABP58225 standard; Protein; 828 AA.
XX
AC ABP58225;
XX
DT 31-MAR-2003 (first entry)
XX
DE Human cell adhesion and extracellular matrix protein 2.
XX
KW Cell adhesion and extracellular matrix protein 2; CADECM-2; human;
KW anti-HIV; virucide; anti-allergic; anti-inflammatory; antianemic;
KW antiparkinsonian; nootropic; anticonvulsant; antiinfertility;
KW antiarteriosclerotic; antiasthmatic; immunosuppressive; antithyroid;
KW cyostatic; hepatotropic; dermatological; antidiabetic; nephrotropic;
KW angiot; thymimetic; neuroprotective; osteopathic; antiarthritic;
KW antiparasitic; antihelminthic; antipsoriatic; uropathic;
KW ophthalmological; antirheumatic; haemostatic; antibacterial;
```

KW protozoacide; fungicide; gynaecological; cadherin; gene therapy.
 XX Homo sapiens.
 XX WO200288322-A2.
 PN 07-NOV-2002.
 XX 01-MAY-2002; 2002WO-US13874.
 XX 02-MAY-2001; 2001US-288290P.
 PR 21-MAY-2001; 2001US-292468P.
 PR 15-JUN-2001; 2001US-298616P.
 PR 28-JUN-2001; 2001US-301672P.
 PR 04-JAN-2002; 2002US-345008P.
 XX (INCY-) INCYTE GENOMICS INC.
 PA Yue H, Lee EA, Duggan BM, Thangavelu K, Honchell CD, Ding L,
 XX Hillman JL, Baughn MR, Kallick DA, Lee S, Warren BA, Xu Y,
 PI Tran UK, Lal PG, Thornton M, Hafalia AJA, Yao MG, Nguyen DB;
 PI Gandhi AR, Khan FA, Walhia NK, Griffin JA, Chinn AM, Elliott VS;
 PI Ramkumar J, Arvizu CS, Forsythe IJ;
 XX WPI; 2003-167112/16.
 DR N-PSDB; AB224579.
 XX New human cell adhesion and extracellular matrix proteins, useful for
 PT diagnosing, treating or preventing autoimmune or inflammatory disorder
 PT (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,
 PT cancer or hepatitis -
 XX Claim 1; Page 140-141; 178pp; English.
 XX The present sequence is the protein sequence of Incyte polypeptide
 CC 1414780CD1 denoted human cell adhesion and extracellular matrix
 CC protein 2 (CADECM-2). The protein is encoded by a clone isolated
 CC from a male cerebellum cDNA library. Homology searches indicate
 CC it to be a cadherin. The invention provides CADECM-1 to -11
 CC polypeptides (see ABP58224-34) and polynucleotides (see AB224578-88),
 CC expression vectors, host cells, antibodies, agonists and antagonists.
 CC These are useful for diagnosing, treating or preventing disorders
 CC associated with aberrant expression of CADECM, particularly cell
 CC proliferative disorders (e.g. arteriosclerosis, atherosclerosis,
 CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria,
 CC polycythaemia vera, psoriasis, primary thrombocytopaenia or
 CC cancer), developmental disorders (e.g. renal tubular acidosis,
 CC anaemia or mental retardation), neurological disorders (e.g.
 CC Alzheimer's disease, Parkinson's disease or epilepsy), reproductive
 CC disorders (e.g. infertility or a disruption in the menstrual cycle),
 CC or autoimmune/inflammatory disorders (e.g. AIDS, allergy, asthma,
 CC autoimmune thyroiditis, contact dermatitis, Crohn's disease,
 CC diabetes mellitus, glomerulonephritis, Goodpasture's syndrome, gout,
 CC Graves' disease, Hashimoto's thyroiditis, irritable bowel syndrome,
 CC multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis,
 CC Reiter's syndrome, rheumatoid arthritis, Sjogren's syndrome, or
 CC uveitis), or viral, bacterial, fungal, parasitic, protozoal or
 CC helminthic infections.
 XX
 SQ Sequence 828 AA;
 Query Match 2.2%; Score 8; DB 24; Length 828;
 Best Local Similarity 100.0%; Pred. No. 89;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 234 AGGGAGGG 241
 |||||
 Db 706 AGGGAGGG 713
 RESULT 39
 AAG91922
 ID AAG91922 standard; Protein; 1004 AA.

XX AAG91922;
 AC
 XX 26-SEP-2001 (first entry)
 DT
 XX C glutamicum protein fragment SEQ ID NO: 5676.
 DE
 XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis.
 KW Corynebacterium glutamicum.
 XX
 OS
 XX EP1108790-A2.
 PN
 XX 20-JUN-2001.
 PD
 XX 18-DEC-2000; 2000EP-0127688.
 PF
 XX 16-DEC-1999; 99JP-0377484.
 PR
 PR 07-APR-2000; 2000JP-0159162.
 PR 03-AUG-2000; 2000JP-0280988.
 XX
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 PA
 XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 PI
 XX WPI; 2001-376931/40.
 DR
 DR N-PSDB; AAH67141.
 XX
 PT Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analysing
 PT expression profile or pattern of a gene and identifying homologous gene
 PT
 XX Claim 17; SEQ ID NO: 5676; 246pp + Sequence Listing; English.
 PS
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and
 CC analysing the expression profile or expression pattern of a gene derived
 CC from Coryneform bacterium, and identifying a homologue of a gene derived
 CC from coryneform bacterium. Coryneform bacteria are useful for producing
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described
 CC in the exemplification of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.
 XX
 SQ Sequence 1004 AA;
 Query Match 2.2%; Score 8; DB 22; Length 1004;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 234 AGGGAGGG 241
 |||||
 Db 338 AGGGAGGG 345
 RESULT 40
 ABP71196
 ID ABP71196 standard; Protein; 1088 AA.
 XX
 AC ABP71196;
 XX
 DT 14-APR-2003 (first entry)
 DE
 XX S. cinamomeus cinnamycin cinM protein.
 XX Cinnamycin; bacterium; cinA; cinM; cinX; cinT; cinH; cinY; antibiotic;
 KW

KW food additive; antibacterial.
 XX
 OS Streptomyces cinamomeus.
 XX
 PN WO200288367-A1.
 XX
 XX WO200288367-A1.
 PD 07-NOV-2002.
 XX
 PF 29-APR-2002; 2002WO-GB01983.
 XX
 PR 27-APR-2001; 2001GB-0010432.
 XX
 XX (PLAN-) PLANT BIOSCIENCE LTD.
 PA (WIDD/) WIDDICK D A.
 XX
 XX Bibb MJ;
 XX
 XX WPI: 2003-111893/10.
 DR N-PSDB; ABZ58812.
 XX
 XX New expression cassettes or genes isolated from Streptomyces
 PT cinamomeus, useful for producing a library of lantibiotic-producing
 PT host cells or lantibiotics, which are useful as food additives and
 PT antibacterial agents -
 XX
 PS Claim 13; Fig 12; 11pp; English.
 XX
 CC The invention relates to expression cassettes or sets of nucleic acids
 CC comprising various open reading frames selected from (a) a cina open
 CC reading frame (orf), a cinM orf, and optionally a cinX orf; or (b) a cina
 CC orf, a cinM orf, a cinT orf, a cinH orf, a cinY orf, and optionally a
 CC cinX orf. The expression cassettes, set of nucleic acids, (set of)
 CC vectors, or methods are useful for producing a library of lantibiotic-
 CC producing host cells or a library of lantibiotics. These are particularly
 CC useful for producing lantibiotic cinamycin or its modified versions. The
 CC lantibiotics are useful as antibiotics having efficacy and utility as
 CC food additives and antibacterial agents against Propionibacterium acnes
 CC and problematic pathogens, e.g. methicillin-resistant Staphylococcus
 CC aureus (which has or is developing resistance to many commonly used
 CC antibiotics), or Streptococcus pneumoniae. Sequences ABP71191-71211
 CC represent the various cinorf proteins encoded by the cinamycin cluster
 CC from S. cinamomeus 40005 as present on the plasmid pDMF79.
 XX
 SQ Sequence 1088 AA;
 Query Match 2.2%; Score 8; DB 24; Length 1088;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 234 AGGGAGGG 241
 Db 1039 AGGGAGGG 1046
 |||||
 |||||
 RESULT 41
 AAU91279
 ID AAU91279 standard; Protein; 1329 AA.
 XX
 AC AAU91279;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Human NOV3a protein.
 XX
 KW Human; NOVX; gene therapy; cardiomyopathy; atherosclerosis;
 KW diabetes; cell signal processing; metabolic pathway modulation;
 KW inflammation; autoimmune disorder; scleroderma; transplantation;
 KW allergy; systemic lupus erythematosus; haemophilia; Alzheimer's disease;
 KW graft versus host disease; Leach-Nyhan syndrome; periodontitis;
 KW pancreatitis; musculoskeletal disorder; Parkinson's disease;
 KW Huntington's disease; behavioural disorder; pain; obesity; wound healing;
 KW neurodegenerative disorder; neuropsychiatric disorder; hypertension;
 KW growth disorder; reproductive disorder; lung disease.

XX Homo sapiens.
 XX
 PN WO200216600-A2.
 XX
 XX 28-FEB-2002.
 PD
 XX
 PF 27-AUG-2001; 2001WO-US26518.
 XX
 PR 25-AUG-2000; 2000US-227800P.
 PR 25-AUG-2000; 2000US-228205P.
 PR 25-AUG-2000; 2000US-228324P.
 PR 30-AUG-2000; 2000US-228997P.
 PR 30-AUG-2000; 2000US-229185P.
 PR 01-SEP-2000; 2000US-229780P.
 PR 01-SEP-2000; 2000US-229848P.
 PR 01-SEP-2000; 2000US-229850P.
 PR 22-JAN-2001; 2001US-263337P.
 PR 31-JAN-2001; 2001US-265518P.
 PR 15-MAR-2001; 2001US-276451P.
 PR 27-MAR-2001; 2001US-279196P.
 PR 24-AUG-2001; 2001US-0393398.
 XX

(CURA-) CURAGEN CORP.

Gerlach V, MacDougall JR, Smithson G, Stone DJ, Ellerman K;
 Spytek KA, Zerhusen BD, Rastelli L, Verney CAM, Patturajan M;
 Tchernev VT, Padigar M, Taupier RJ;

WPI: 2002-292064/33.
 N-PSDB; ABK5563.

New isolated cytoplasmic, nuclear, membrane bound and secreted
 polypeptides, termed NOVX, useful for treating inflammation, autoimmune
 disorders, haemophilia, Leach-Nyhan syndrome, pancreatitis,
 musculoskeletal disorders -

Claim 1; Page 32; 245pp; English.

The invention relates to an isolated cytoplasmic, nuclear, membrane bound
 or secreted polypeptide, designated NOVX (actually NOV1, 2a, 2b, 3a, 3b,
 4, 5a, 5b, 5c, 5d, 5f, 5g, 5h, 5i, 6, 7 and 8), a variant of NOVX, a
 mature form, or a variant of the mature form of NOVX. Also included
 are a polynucleotide encoding NOVX (or its complement), a vector
 comprising the polynucleotide, a cell comprising the vector, an
 anti-NOVX antibody, determining the presence of NOVX in a sample
 using the antibody, determining the presence of NOVX polynucleotide in
 a sample using a probe which binds to NOVX polynucleotide, identifying a
 an agent which binds to NOVX (including modulators of NOVX). NOVX, the
 polynucleotide and the antibody are useful for diagnosing, treating
 or preventing a NOVX-associated disorder selected from cardiomyopathy,
 atherosclerosis, diabetes, a disorder related to cell signal processing,
 and metabolic pathway modulation, inflammation, autoimmune disorders,
 scleroderma, transplantation, allergies, systemic lupus erythematosus,
 haemophilia, graft versus host disease, Alzheimer's disease, stroke,
 Leach-Nyhan syndrome, periodontitis, pancreatitis, musculoskeletal
 disorders, Parkinson's disease, Huntington's disease, behavioural
 disorders, pain, neurodegenerative and neuropsychiatric disorders,
 hypertension, wound healing, obesity, growth and reproductive
 disorders, lung diseases and many other diseases and disorders listed in
 the specification. NOVX, the polynucleotide and the antibody are useful
 in screening assays, detection assays (e.g., chromosomal mapping, tissue
 typing, forensic biology), predictive medicine (e.g., diagnostic assays,
 prognostic assays, monitoring clinical trials and pharmacogenomic), and
 in methods of treatment (e.g., therapeutic and prophylactic). NOVX is
 useful as immunogen to produce antibodies immunospecific for NOVX, as
 vaccines to screen for potential agonist and antagonist compounds; and
 as bait protein in a two-hybrid or three-hybrid assay. The polynucleotide
 is useful in gene therapy, to express NOVX, to detect NOVX mRNA
 or a genetic lesion in a NOVX gene, and to modulate NOVX activity. The
 vector is useful for producing non-human transgenic animals. The antibody
 is useful for isolating, and purifying NOVX and to monitor protein levels
 in tissue as part of a clinical testing procedure. The present sequence

CC represents a NOVX protein.

XX Sequence 1329 AA;

Query Match 2.2%; Score 8; DB 23; Length 1329;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241

|||||||

Db 745 AGGGAGGG 752

RESULT 42

ABB60138

ID ABB60138 standard; Protein; 1385 AA.

XX

AC ABB60138;

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster polypeptide SEQ ID NO 7206.

XX

KW Drosophila; developmental biology; cell signalling; insecticide;

KW Pharmaceutical.

XX

OS Drosophila melanogaster.

XX

PN WO200171042-A2.

XX

PD 27-SEP-2001.

XX

PF 23-MAR-2001; 2001WO-US09231.

XX

PR 23-MAR-2000; 2000US-191637P.

XX

PR 11-JUL-2000; 2000US-0614150.

XX

PA (PEKE) PE CORP NY.

XX

PI Venter JC, Adams M, Li PWD, Myers EW;

XX

WPI; 2001-656860/75.

XX

DR N-PSDB; ABL04241.

XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more

genes from Drosophila and for elucidating cell signalling and cell-cell

interactions -

XX

PS Disclosure; SEQ ID NO 7206; 2lpp + Sequence Listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent

capable of detecting 1000 or more genes from Drosophila. The invention is

useful in developmental biology and in elucidating cell signalling and

cell-cell interactions in higher eukaryotes for the development of

insecticides, therapeutics and pharmaceutical drugs. The invention

discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

sequences (ABL01840-ABL16175) and the encoded proteins

(ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed

specification, but was obtained in electronic format directly from WIPO

at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 1385 AA;

Query Match

Best Local Similarity 2.2%; Score 8; DB 22; Length 1385;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241

|||||||

Db 344 AGGGAGGG 351

RESULT 43

ABB65356

ID ABB65356 standard; Protein; 1417 AA.

XX

AC ABB65356;

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster polypeptide SEQ ID NO 22860.

XX

KW Drosophila; developmental biology; cell signalling; insecticide;

KW Pharmaceutical.

XX

OS Drosophila melanogaster.

XX

PN WO200171042-A2.

XX

PD 27-SEP-2001.

XX

PF 23-MAR-2001; 2001WO-US09231.

XX

PR 23-MAR-2000; 2000US-191637P.

XX

PR 11-JUL-2000; 2000US-0614150.

XX

PA (PEKE) PE CORP NY.

XX

PI Venter JC, Adams M, Li PWD, Myers EW;

XX

WPI; 2001-656860/75.

XX

DR N-PSDB; ABL09459.

XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more

genes from Drosophila and for elucidating cell signalling and cell-cell

interactions -

XX

PS Disclosure; SEQ ID NO 22860; 2lpp + Sequence Listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent

capable of detecting 1000 or more genes from Drosophila. The invention is

useful in developmental biology and in elucidating cell signalling and

cell-cell interactions in higher eukaryotes for the development of

insecticides, therapeutics and pharmaceutical drugs. The invention

discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

sequences (ABL01840-ABL16175) and the encoded proteins

(ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed

specification, but was obtained in electronic format directly from WIPO

at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 1417 AA;

Query Match

Best Local Similarity 2.2%; Score 8; DB 22; Length 1417;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241

|||||||

Db 719 AGGGAGGG 726

RESULT 44

ABB61677

ID ABB61677 standard; Protein; 2075 AA.

XX

AC ABB61677;

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster polypeptide SEQ ID NO 11823.

XX

KW Drosophila; developmental biology; cell signalling; insecticide;

KW Pharmaceutical.

XX

CC Alzheimer's disease, Down's syndrome, oedema, lupus (SLE), vasculitis,
 CC glomerulonephritis, haemolytic anaemia, thrombocytopaenia, arthritis,
 CC meningitis, muscular disorders, dementia, neurological diseases, tubercu-
 CC sclerosis, male infertility, hypercalcaemia, blood pressure disorders,
 CC osteoporosis, pathogenic infections, hypercholesterolaemia, obesity or
 CC autoimmunity. The present sequence is a peptide encoded by a
 CC polymorphism-containing oligonucleotide fragment of the invention.

XX Sequence 9 AA;

Query Match 1.9%; Score 7; DB 22; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05; Mismatches 0; Indels 0; Gaps 0;
 Matches 7; Conservative 0;

QY 236 GGAGGGP 242
 DB 1 GGAGGGP 7
 |||||

RESULT 47
 AAM00521
 ID AAM00521 standard; Peptide; 10 AA.

XX AC AAM00521;
 XX DT 01-OCT-2001 (first entry)
 XX DE Human kinase fragment SEQ ID NO: 1069.
 XX KW Human; single nucleotide polymorphism; SNP; paternity test;
 XX KW forensic test; aberrant protein expression.
 XX OS Homo sapiens.

XX PN WO200151670-A2.
 XX PD 19-JUL-2001.

XX PF 05-JAN-2001; 2001WO-US00322.
 XX PR 07-JAN-2000; 2000US-0174962.
 XX PA (CURA-) CURAGEN CORP.

XX PI Shimkets RA, Leach MD;
 XX DR WPI; 2001-451871/48.
 XX DR N-PSDB; AAH89638.

XX PT Isolated human polynucleotides containing single nucleotide
 XX PT polymorphisms, useful for the treatment and diagnosis of e.g. cancer,
 XX PT infection and diabetes -

XX PS Disclosure; Page 410; 475pp; English.

XX CC The present invention relates to human nucleic acids containing single
 CC nucleotide polymorphisms (SNPs). These can be used in forensic and
 CC paternity tests, and to aid in the treatment of diseases associated with
 CC aberrant protein expression, including cancer, amyloidosis, diabetes,
 CC Alzheimer's disease, Down's syndrome, oedema, lupus (SLE), vasculitis,
 CC glomerulonephritis, haemolytic anaemia, thrombocytopaenia, arthritis,
 CC meningitis, muscular disorders, dementia, neurological diseases, tubercu-
 CC sclerosis, male infertility, hypercalcaemia, blood pressure disorders,
 CC osteoporosis, pathogenic infections, hypercholesterolaemia, obesity or
 CC autoimmunity. The present sequence is a peptide encoded by a
 CC polymorphism-containing oligonucleotide fragment of the invention.

XX Sequence 10 AA;

Query Match 1.9%; Score 7; DB 22; Length 10;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 GGAGGGP 242
 DB 2 GGAGGGP 8
 |||||

RESULT 48
 AAM00522
 ID AAM00522 standard; Peptide; 10 AA.

XX AC AAM00522;
 XX DT 01-OCT-2001 (first entry)
 XX DE Human kinase fragment SEQ ID NO: 1070.
 XX KW Human; single nucleotide polymorphism; SNP; paternity test;
 XX KW forensic test; aberrant protein expression.
 XX OS Homo sapiens.

XX PN WO200151670-A2.
 XX PD 19-JUL-2001.

XX PF 05-JAN-2001; 2001WO-US00322.
 XX PR 07-JAN-2000; 2000US-0174962.
 XX PA (CURA-) CURAGEN CORP.

XX PI Shimkets RA, Leach MD;
 XX DR WPI; 2001-451871/48.
 XX DR N-PSDB; AAH89639.

XX PT Isolated human polynucleotides containing single nucleotide
 XX PT polymorphisms, useful for the treatment and diagnosis of e.g. cancer,
 XX PT infection and diabetes -

XX PS Disclosure; Page 410; 475pp; English.

XX CC The present invention relates to human nucleic acids containing single
 CC nucleotide polymorphisms (SNPs). These can be used in forensic and
 CC paternity tests, and to aid in the treatment of diseases associated with
 CC aberrant protein expression, including cancer, amyloidosis, diabetes,
 CC Alzheimer's disease, Down's syndrome, oedema, lupus (SLE), vasculitis,
 CC glomerulonephritis, haemolytic anaemia, thrombocytopaenia, arthritis,
 CC meningitis, muscular disorders, dementia, neurological diseases, tubercu-
 CC sclerosis, male infertility, hypercalcaemia, blood pressure disorders,
 CC osteoporosis, pathogenic infections, hypercholesterolaemia, obesity or
 CC autoimmunity. The present sequence is a peptide encoded by a
 CC polymorphism-containing oligonucleotide fragment of the invention.

XX Sequence 10 AA;

Query Match 1.9%; Score 7; DB 22; Length 10;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 GGAGGGP 242
 DB 2 GGAGGGP 8
 |||||

RESULT 49
 AAG97623
 ID AAG97623 standard; Peptide; 10 AA.

XX AC AAG97623;
 XX DT 18-SEP-2001 (first entry)
 XX DE Human complementary peptide, SEQ ID NO: 3818.


```
Db          |||||
            2 GGGAGGG 8

RESULT 52
AAG97638
ID AAG97638 standard; Peptide; 10 AA.
XX
AC AAG97638;
XX
DT 18-SEP-2001 (first entry)
XX
DE Human complementary peptide, SEQ ID NO: 3833.
XX
KW Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS Homo sapiens.
XX
PN WO200142277-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-GB04776.
XX
PR 13-DEC-1999; 99GB-0029464.
XX
PA (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
XX
WPI; 2001-408419/43.
XX
A set of peptide ligands consisting of specific complementary peptides
to proteins encoded by genes of the human genome, useful in an assay
for screening and identifying of one or more novel peptides which are
drug candidates or pro-drugs -
XX
PS Example 6; Page 593; 646pp; English.
XX
CC The invention relates to a set of complementary peptide ligands
generated from the human genome. The complementary peptides
interact with their relevant target proteins encoded in the human
genome. They can be used as reagents in drug discovery and as lead
ligands to facilitate drug design and development. The present
sequence is a complementary peptide provided in the specification.
XX
SQ Sequence 10 AA;

Query Match          1.9%; Score 7; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
Db          |||||
            3 GGGAGGG 9

RESULT 53
AAG83526
ID AAG83526 standard; Peptide; 10 AA.
XX
AC AAG83526;
XX
DT 11-SEP-2001 (first entry)
XX
DE Arabidopsis thaliana peptide ligand #166.
XX
KW Plant; peptide pesticide; peptide herbicide; agricultural research.
XX
OS Arabidopsis thaliana.
XX
PN WO200142279-A2.
XX

Query Match          1.9%; Score 7; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
Db          |||||
            3 GGGAGGG 9

RESULT 54
AAG84044
ID AAG84044 standard; Peptide; 10 AA.
XX
AC AAG84044;
XX
DT 11-SEP-2001 (first entry)
XX
DE Arabidopsis thaliana peptide ligand #684.
XX
KW Plant; peptide pesticide; peptide herbicide; agricultural research.
XX
OS Arabidopsis thaliana.
XX
PN WO200142279-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-GB04781.
XX
PR 13-DEC-1999; 99GB-0029469.
XX
PA (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
XX
WPI; 2001-381629/40.
XX
A set of peptide ligands for agricultural research and development of
therapeutic agents comprise specific complementary peptides to proteins
encoded by genes of plant genomes -
XX
PS Example 4; Page 126; 201pp; English.
XX
CC The present invention relates to a set of peptide ligands consisting of
specific complementary peptides to proteins encoded by genes of plant
```

```
PD 14-JUN-2001.
XX
XX 13-DEC-2000; 2000WO-GB04781.
XX
PR 13-DEC-1999; 99GB-0029469.
XX
PA (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
XX
WPI; 2001-381629/40.
XX
A set of peptide ligands for agricultural research and development of
therapeutic agents comprise specific complementary peptides to proteins
encoded by genes of plant genomes -
XX
PS Example 4; Page 54; 201pp; English.
XX
CC The present invention relates to a set of peptide ligands consisting of
specific complementary peptides to proteins encoded by genes of plant
genomes. The present sequence is one such peptide from Arabidopsis
thaliana. The peptides of the present invention are useful in an assay to
identify a peptide, especially a peptide pesticide or herbicide. The
peptides are also useful for tools for agricultural research and
development.
XX
SQ Sequence 10 AA;

Query Match          1.9%; Score 7; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
Db          |||||
            3 GGGAGGG 9

RESULT 54
AAG84044
ID AAG84044 standard; Peptide; 10 AA.
XX
AC AAG84044;
XX
DT 11-SEP-2001 (first entry)
XX
DE Arabidopsis thaliana peptide ligand #684.
XX
KW Plant; peptide pesticide; peptide herbicide; agricultural research.
XX
OS Arabidopsis thaliana.
XX
PN WO200142279-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-GB04781.
XX
PR 13-DEC-1999; 99GB-0029469.
XX
PA (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
XX
WPI; 2001-381629/40.
XX
A set of peptide ligands for agricultural research and development of
therapeutic agents comprise specific complementary peptides to proteins
encoded by genes of plant genomes -
XX
PS Example 4; Page 126; 201pp; English.
XX
CC The present invention relates to a set of peptide ligands consisting of
specific complementary peptides to proteins encoded by genes of plant
```

CC genomes. The present sequence is one such peptide from Arabidopsis
 CC thaliana. The peptides of the present invention are useful in an assay to
 CC identify a peptide, especially a peptide pesticide or herbicide. The
 CC peptides are also useful for tools for agricultural research and
 CC development.

XX SQ Sequence 10 AA;

Query Match 1.9%; Score 7; DB 22; Length 10;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 GGGAGGG 241
 |||||
 Db 1 GGGAGGG 7

RESULT 55

AAR92855
 ID AAR92855 standard; peptide; 12 AA.

XX AC AAR92855;

XX DT 03-OCT-1996 (first entry)

XX DE Cell adhesion modulatory peptide AP9.

XX KW Intercellular adhesion; stimulation; inhibition; skin graft;
 KW synthetic blood vessel; coating; endothelial cell; epidermal cell;
 KW chemotactic attractor; wound healing; organ transplantation;
 KW thrombosis; arteriosclerosis; cancer metastases.

XX OS Synthetic.

XX PN DE4430601-A1.

XX FD 29-FEB-1996.

XX PF 22-AUG-1994; 94DE-4430601.

XX PR 22-AUG-1994; 94DE-4430601.

XX FA (BEIE) BEIERSDORF AG.

XX PI Doerschner A, Eichner W, Kock K, Mielke H;

XX DR WPI; 1996-130242/14.

XX PT Peptide(s) that stimulate or inhibit cell to cell adhesion - used
 PT e.g. to coat synthetic blood vessels with endothelial cells, to
 PT prepare, or increase growth of skin grafts, to prevent thrombosis
 PT etc.

XX PS Claim 12; Page 14; 18pp; German.

XX CC The present peptide is a specifically claimed example of a peptide
 CC which contains the highly generic sequence AA5-AA4-AA3-AA2-AA1-(AAAX)n
 CC where AA5 is Glu, Ser, Asp or Asn; AA4 is Leu or Ser; AA3 is Leu, Ile,
 CC Phe or Gly; AA2 is Asp, Leu, Asn or Ser; AA1 is Gly, Pro or Asp; AAX
 CC is any amino acid and n = 0 or 1. When two or more such peptides are
 CC attached to a carrier, the product can be used for stimulating
 CC adhesion of eukaryotic cells in vitro. Particular applications include
 CC coating synthetic blood vessels with endothelial cells, preparing skin
 CC grafts using epithelial cells or stimulating wound healing. When a
 CC single peptide is used it may inhibit intercellular adhesion, making
 CC it useful for preventing thrombosis or arteriosclerosis or to suppress
 CC cancer metastases. The peptides can also be used as chemotactic
 CC attractors and for detecting/quantifying cell-cell adhesion in vitro.

XX SQ Sequence 12 AA;

Query Match 1.9%; Score 7; DB 17; Length 12;
 Best Local Similarity 100.0%; Pred. No. 16;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 235 GGGAGGG 241
 |||||
 Db 1 GGGAGGG 7

RESULT 56

AAR24435
 ID AAR24435 standard; peptide; 14 AA.

XX AC AAR24435;

XX DT 21-NOV-1992 (first entry)

XX DE Sequence of synthetic Epstein-Barr virus (EBV)-encoded nuclear
 DE antigen (EBNA) designated D3.

XX KW Herpes virus; Epstein-Barr virus; infectious mononucleosis;
 KW Burkett's lymphoma; nasopharyngeal carcinoma; rheumatoid arthritis.
 XX OS Synthetic.

XX PN US5116725-A.

XX PD 26-MAY-1992.

XX PF 04-NOV-1987; 87US-0117241.

XX PR 08-AUG-1984; 84US-0638726.

XX PR 24-MAR-1987; 87US-0029860.

XX PR 04-NOV-1987; 87US-0117241.

XX PA (SCRI) SCRIPPS CLINIC & RES FOUND.

XX PI Carson DA, Geltosky JE, Houghten R, Rhodes G, Smith RS;

XX PI Vaughan JH;

XX DR WPI; 1992-199513/24.

XX PT Determining state of Epstein-Barr virus infection - by
 PT determining ratio of IgG to IgM antibodies to synthetic peptides
 PT in patient

XX PS Claim Iaviii; Column 54; 40pp; English.

XX CC EBNA-1 is the earliest EBV-associated antigen that can be detected
 CC after infection. EBNA has been detected in the nucleus of latently-
 CC infected growth-transformed B lymphocytes. EBNA has also been
 CC detected in the nuclei of African Burkitt tumor lymphoblasts and
 CC anaplastic nasopharyngeal carcinoma cells. The size of the EBNA-1
 CC protein is correlated with the variation of the length of the IR-3
 CC region of EBV-DNA. The IR-3 region encodes a repeating glycine-
 CC alanine sequence that has been characterised to be the major epitope
 CC of the EBNA-1 protein. The peptides of the invention are synthetic,
 CC random copolymeric polypeptides capable of inducing the production of
 CC antibodies that immunoreact with EBNA and of immunoreacting with
 CC human antibodies induced by EBNA. The sequences were chosen to
 CC represent different areas from within and just outside the proposed
 CC IR-3 region of the EBNA.

XX SQ Sequence 14 AA;

Query Match 1.9%; Score 7; DB 13; Length 14;
 Best Local Similarity 100.0%; Pred. No. 18;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 240
 |||||
 Db 6 AGGAGGG 12

RESULT 57

AAR25057
 ID AAR25057 standard; Protein; 14 AA.
 AC
 AC AAR25057;
 XX
 XX
 DT 25-MAR-2003 (updated)
 DT 09-DEC-1992 (first entry)
 FT
 FT Synthetic random copolymer peptide #2.
 XX
 DE EBNA; antibody; IM; EBV; cytomegalovirus.
 XX
 KW EBNA; antibody; IM; EBV; cytomegalovirus.
 XX
 OS Synthetic.
 XX
 PN US5122448-A.
 XX
 XX 16-JUN-1992.
 PD
 XX 11-JAN-1990; 90US-0463505.
 PF
 XX 08-AUG-1984; 84US-0638726.
 PR
 PR 24-MAR-1987; 87US-0029860.
 PR
 PR 04-NOV-1987; 87US-0117241.
 PR
 PR 11-JAN-1990; 90US-0463505.
 XX
 XX (SCRI) SCRIPPS CLINIC & RES FOUND.
 PA
 XX Carson DA, Houghten R, Rhodes G, Vaughan JH;
 PI WPI; 1992-226077/27.
 XX
 XX Assay for anti-Epstein Barr virus nuclear antigen antibodies -
 PT using random copolymer polypeptide contg. specified aminoacid
 PT residues, for diagnosis of infectious mononucleosis and
 PT nasopharyngeal carcinoma
 XX
 XX Claim 1; Page 54; 41pp; English.
 PS
 CC The sequences given in AAR25050-7 are random copolymer peptides which
 CC can be used in diagnostic methods and systems relating to Epstein-
 CC Barr virus nuclear antigen (EBNA). These peptides are capable of
 CC inducing the production of antibodies which immunoreact with EBNA,
 CC and of immunoreacting with human antibodies induced by EBNA. The
 CC peptides can be fixed to a solid matrix, giving a solid
 CC immunoreactant in a diagnostic kit. The assay developed using
 CC these peptides has been found to be clinically reliable in
 CC detecting infectious mononucleosis (IM) caused by Epstein-Barr virus
 CC (EBV) as well as IM induced by cytomegalovirus, and also in
 CC detecting nasopharyngeal carcinoma, another disease in which EBV
 CC has been implicated.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 CC
 XX Sequence 14 AA;
 SQ
 Query Match 1.9%; Score 7; DB 13; Length 14;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 234 AGGGAGG 240
 Db 6 AGGGAGG 12
 |||||
 RESULT 58
 AAP60478
 ID AAP60478 standard; protein; 15 AA.
 XX
 XX AAP60478;
 AC
 AC AAP60478;
 XX
 DT 25-MAR-2003 (updated)
 DT 28-JUL-1991 (first entry)
 FT
 FT Epstein-Barr virus nuclear antigen (EBNA) pentapeptide.
 XX
 DE

XX Epstein-Barr virus; nuclear antigen; passive immunization; diagnosis.
 KW
 XX Synthetic.
 OS
 XX
 XX Key Location/Qualifiers
 FH 6..10
 FT Region /label= EBNA determinant
 FT
 XX WO8601210-A.
 PN
 XX 27-FEB-1986.
 PD
 XX 02-AUG-1985; 85WO-US01484.
 PF
 XX 08-AUG-1984; 84US-0638726.
 PR
 XX (SCRI) SCRIPPS CLINIC & RES FOUND.
 PA
 XX Vaughan JH, Carbon DA, Rhodes G, Houghten R;
 PI WPI; 1986-068962/10.
 DR
 XX New synthetic copolymer polypeptide(s) - useful for preventing
 PT and diagnosing diseases involving Epstein-Barr virus.
 PT
 XX Claim 4(vi); Page 73; 84pp; English.
 PS
 CC The peptide includes the EBNA pentapeptide which is an
 CC immunogen for the production of an antibody in a host animal.
 CC Epstein-Barr virus- and EBNA-associated diseases may be
 CC diagnosed and prevented.
 CC See also AAP60473-P60483
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC
 XX Sequence 15 AA;
 SQ
 Query Match 1.9%; Score 7; DB 7; Length 15;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 234 AGGGAGG 240
 Db 7 AGGGAGG 13
 |||||
 RESULT 59
 ABR38299
 ID ABR38299 standard; Peptide; 15 AA.
 XX
 XX ABR38299;
 AC
 XX 19-MAY-2003 (first entry)
 DT
 XX Human cancer-related protein 187P3F2 HLA peptide #1430.
 DE
 XX Human; cytostatic; vaccine; cancer; immune response; HLA;
 KW human leukocyte antigen.
 KW
 XX Homo sapiens.
 OS
 XX WO200283921-A2.
 PN
 XX 24-OCT-2002.
 PD
 XX 10-APR-2002; 2002WO-US11654.
 PF
 XX 10-APR-2001; 2001US-282739P.
 PR
 PR 10-APR-2001; 2001US-283112P.
 PR
 PR 25-APR-2001; 2001US-286630P.
 XX
 XX (AGEN-) AGENSYS INC.
 PA

PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
 PI Morrison K, Morrison RK, Raitano AB;
 DR WPI; 2003-075555/07.
 XX
 XX New composition comprising a substance that modulates the structure of
 PT proteins and polynucleotides, useful for therapeutic, prognostic and
 PT diagnostic reagents for eliciting cellular or humoral immune response
 PT in cancer patients -
 XX
 XX Claim 13; Page 617; 1021pp; English.
 XX
 CC The present invention relates to novel human cancer-related genes and
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.
 XX
 SQ Sequence 15 AA;
 Query Match 1.9%; Score 7; DB 24; Length 15;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 235 GGGAGGG 241
 Db |||||
 9 GGGAGGG 15
 RESULT 60
 ABR38300
 ID ABR38300 standard; Peptide; 15 AA.
 XX
 AC ABR38300;
 XX
 DT 19-MAY-2003 (first entry)
 XX
 DE Human cancer-related protein 187P3F2 HLA peptide #1431.
 KW Human; cytostatic; vaccine; cancer; immune response; HLA;
 KW human leukocyte antigen.
 XX
 OS Homo sapiens.
 XX
 PN WO200283921-A2.
 XX
 PD 24-OCT-2002.
 XX
 PF 10-APR-2002; 2002WO-US11654.
 XX
 PR 10-APR-2001; 2001US-282739P.
 PR 10-APR-2001; 2001US-283112P.
 PR 25-APR-2001; 2001US-286630P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
 PI Morrison K, Morrison RK, Raitano AB;
 DR WPI; 2003-075555/07.
 XX
 XX New composition comprising a substance that modulates the structure of
 PT proteins and polynucleotides, useful for therapeutic, prognostic and
 PT diagnostic reagents for eliciting cellular or humoral immune response
 PT in cancer patients -
 XX

PS Claim 13; Page 617; 1021pp; English.
 XX
 CC The present invention relates to novel human cancer-related genes and
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.
 XX
 SQ Sequence 15 AA;
 Query Match 1.9%; Score 7; DB 24; Length 15;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 235 GGGAGGG 241
 Db |||||
 2 GGGAGGG 8
 RESULT 61
 ABR38302
 ID ABR38302 standard; Peptide; 15 AA.
 XX
 AC ABR38302;
 XX
 DT 19-MAY-2003 (first entry)
 XX
 DE Human cancer-related protein 187P3F2 HLA peptide #1433.
 KW Human; cytostatic; vaccine; cancer; immune response; HLA;
 KW human leukocyte antigen.
 XX
 OS Homo sapiens.
 XX
 PN WO200283921-A2.
 XX
 PD 24-OCT-2002.
 XX
 PF 10-APR-2002; 2002WO-US11654.
 XX
 PR 10-APR-2001; 2001US-282739P.
 PR 10-APR-2001; 2001US-283112P.
 PR 25-APR-2001; 2001US-286630P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
 PI Morrison K, Morrison RK, Raitano AB;
 DR WPI; 2003-075555/07.
 XX
 XX New composition comprising a substance that modulates the structure of
 PT proteins and polynucleotides, useful for therapeutic, prognostic and
 PT diagnostic reagents for eliciting cellular or humoral immune response
 PT in cancer patients -
 XX
 XX Claim 13; Page 617; 1021pp; English.
 XX
 CC The present invention relates to novel human cancer-related genes and
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.
 XX

CC as therapeutic agents. The proteins and peptides are useful as
CC therapeutic, prognostic and diagnostic reagents for cancer. The present
CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
CC from the invention.

XX SQ Sequence 15 AA;

Query Match 1.9%; Score 7; DB 24; Length 15;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
|||||||
Db 9 GGGAGGG 15

RESULT 62

ABR38314
ID ABR38314 standard; Peptide; 15 AA.

XX AC ABR38314;

XX XX 19-MAY-2003 (first entry)

XX DE Human cancer-related protein 187P3F2 HLA peptide #1445.

XX XX Human; cytostatic; vaccine; cancer; immune response; HLA;
KW human leukocyte antigen.

XX OS Homo sapiens.

XX PN WO200283921-A2.

XX PD 24-OCT-2002.

XX PF 10-APR-2002; 2002WO-US11654.

XX PR 10-APR-2001; 2001US-282739P.

XX PR 10-APR-2001; 2001US-283112P.

XX PR 25-APR-2001; 2001US-286630P.

XX PA (AGEN-) AGENSYS INC.

XX PI Jakobovits A, Challita-Bid PM, Faris M, Ge W, Hubert RS;

XX PI Morrison K, Morrison RK, Raitano AB;

XX XX WPI; 2003-075555/07.

XX PT New composition comprising a substance that modulates the structure of
PT proteins and polynucleotides, useful for therapeutic, prognostic and
PT diagnostic reagents for eliciting cellular or humoral immune response
PT in cancer patients -

XX PS Claim 13; Page 617; 1021pp; English.

XX CC The present invention relates to novel human cancer-related genes and
CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC proteins are useful for eliciting a humoral or cellular immune response.
CC The genes are useful as probes and primers for the amplification and/or
CC detection of genes, mRNAs or their fragments, as reagents for the
CC diagnosis and/or prognosis of cancer, as coding sequences capable of
CC directing the expression of the protein, as tools for modulating or
CC inhibiting the expression of genes and/or translation of transcripts, and
CC as therapeutic agents. The proteins and peptides are useful as
CC therapeutic, prognostic and diagnostic reagents for cancer. The present
CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
CC from the invention.

XX SQ Sequence 15 AA;

Query Match 1.9%; Score 7; DB 24; Length 15;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
|||||||
Db 5 GGGAGGG 11

RESULT 63

ABR38327

ID ABR38327 standard; Peptide; 15 AA.

XX AC ABR38327;

XX DT 19-MAY-2003 (first entry)

XX DE Human cancer-related protein 187P3F2 HLA peptide #1458.

XX KW Human; cytostatic; vaccine; cancer; immune response; HLA;
KW human leukocyte antigen.

XX OS Homo sapiens.

XX PN WO200283921-A2.

XX PD 24-OCT-2002.

XX PF 10-APR-2002; 2002WO-US11654.

XX PR 10-APR-2001; 2001US-282739P.

XX PR 10-APR-2001; 2001US-283112P.

XX PR 25-APR-2001; 2001US-286630P.

XX PA (AGEN-) AGENSYS INC.

XX PI Jakobovits A, Challita-Bid PM, Faris M, Ge W, Hubert RS;

XX PI Morrison K, Morrison RK, Raitano AB;

XX DR WPI; 2003-075555/07.

XX PT New composition comprising a substance that modulates the structure of
PT proteins and polynucleotides, useful for therapeutic, prognostic and
PT diagnostic reagents for eliciting cellular or humoral immune response
PT in cancer patients -

XX PS Claim 13; Page 617; 1021pp; English.

XX CC The present invention relates to novel human cancer-related genes and
CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC proteins are useful for eliciting a humoral or cellular immune response.
CC The genes are useful as probes and primers for the amplification and/or
CC detection of genes, mRNAs or their fragments, as reagents for the
CC diagnosis and/or prognosis of cancer, as coding sequences capable of
CC directing the expression of the protein, as tools for modulating or
CC inhibiting the expression of genes and/or translation of transcripts, and
CC as therapeutic agents. The proteins and peptides are useful as
CC therapeutic, prognostic and diagnostic reagents for cancer. The present
CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
CC from the invention.

XX SQ Sequence 15 AA;

Query Match 1.9%; Score 7; DB 24; Length 15;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
|||||||
Db 3 GGGAGGG 9

RESULT 64

ABR38328

ID ABR38328 standard; Peptide; 15 AA.

XX XX

AC ABR38328;
 XX 19-MAY-2003 (first entry)
 DT Human cancer-related protein 187P3F2 HLA peptide #1459.
 DE Human; cytostatic; vaccine; cancer; immune response; HLA;
 XX human leukocyte antigen.
 KW Homo sapiens.
 XX WO200283921-A2.
 PN 24-OCT-2002.
 XX 10-APR-2002; 2002WO-US11654.
 XX 10-APR-2001; 2001US-282739P.
 PR 10-APR-2001; 2001US-283112P.
 PR 25-APR-2001; 2001US-286630P.
 XX (AGEN-) AGENSYS INC.
 PA Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
 PI Morrison K, Morrison RK, Raitano AB;
 PI WPI; 2003-075555/07.
 DR New composition comprising a substance that modulates the structure of
 XX proteins and polynucleotides, useful for therapeutic, prognostic and
 PT diagnostic reagents for eliciting cellular or humoral immune response
 PT in cancer patients
 PT Claim 13; Page 617; 1021pp; English.
 PS The present invention relates to novel human cancer-related genes and
 XX proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC diagnostic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.
 XX SQ Sequence 15 AA;
 Query Match 1.9%; Score 7; DB 24; Length 15;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 235 GGGAGGG 241
 Db |||||
 2 GGGAGGG 8
 RESULT 65
 ABR38464
 ID ABR38464 standard; Peptide; 15 AA.
 XX ABR38464;
 AC 19-MAY-2003 (first entry)
 DT Human cancer-related protein 187P3F2 HLA peptide #1595.
 DE Human; cytostatic; vaccine; cancer; immune response; HLA;
 XX human leukocyte antigen.
 KW Homo sapiens.
 OS

XX WO200283921-A2.
 XX 24-OCT-2002.
 XX 10-APR-2002; 2002WO-US11654.
 XX 10-APR-2001; 2001US-282739P.
 PR 10-APR-2001; 2001US-283112P.
 PR 25-APR-2001; 2001US-286630P.
 XX (AGEN-) AGENSYS INC.
 PA Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
 PI Morrison K, Morrison RK, Raitano AB;
 PI WPI; 2003-075555/07.
 DR New composition comprising a substance that modulates the structure of
 XX proteins and polynucleotides, useful for therapeutic, prognostic and
 PT diagnostic reagents for eliciting cellular or humoral immune response
 PT in cancer patients
 PT Claim 13; Page 619; 1021pp; English.
 PS The present invention relates to novel human cancer-related genes and
 XX proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC diagnostic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.
 XX SQ Sequence 15 AA;
 Query Match 1.9%; Score 7; DB 24; Length 15;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 235 GGGAGGG 241
 Db |||||
 2 GGGAGGG 8
 RESULT 66
 AAR24433
 ID AAR24433 standard; peptide; 16 AA.
 XX AAR24433;
 AC 21-NOV-1992 (first entry)
 DT Sequence of synthetic Epstein-Barr virus (EBV)-encoded nuclear
 DE antigen (EBNA) designated F15.
 DE Herpes virus; Epstein-Barr virus; infectious mononucleosis;
 XX Burkett's lymphoma; nasopharyngeal carcinoma; rheumatoid arthritis.
 KW Synthetic.
 XX US5116725-A.
 PN 26-MAY-1992.
 XX 04-NOV-1987; 87US-0117241.
 PF 08-AUG-1984; 84US-0638726.
 PR 24-MAR-1987; 87US-0029860.
 PR

CC infected growth-transformed B lymphocytes. EBNA has also been
 CC detected in the nuclei of African Burkitt tumor lymphoblasts and
 CC anaplastic nasopharyngeal carcinoma cells. The size of the EBNA-1
 CC protein is correlated with the variation of the length of the IR-3
 CC region of EBV-DNA. The IR-3 region encodes a repeating glycine-
 CC alanine sequence that has been characterised to be the major epitope
 CC of the EBNA-1 protein. The peptides of the invention are synthetic,
 CC random copolymeric polypeptides capable of inducing the production of
 CC antibodies that immunoreact with EBNA and of immunoreacting with
 CC human antibodies induced by EBNA. The sequences were chosen to
 CC represent different areas from within and just outside the proposed
 CC IR-3 region of the EBNA.

XX Sequence 17 AA;

Query Match 1.9%; Score 7; DB 13; Length 17;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGG 240
 |||||
 Db 9 AGGGAGG 15

RESULT 69

AAAR25056
 ID AAR25056 standard; Protein; 17 AA.

XX AAR25056;

DT 25-MAR-2003 (updated)

DT 09-DEC-1992 (first entry)

XX Synthetic random copolymer peptide #1.

DE EBNA; antibody; IM; EBV; cytomegalovirus.

XX Synthetic.

XX US5122448-A.

PD 16-JUN-1992.

XX 11-JAN-1990; 90US-0463505.

XX 08-AUG-1984; 84US-0638726.

PR 24-MAR-1987; 87US-0029860.

PR 04-NOV-1987; 87US-0117241.

PR 11-JAN-1990; 90US-0463505.

XX (SCRI) SCRIPPS CLINIC & RES FOUND.

XX Carson DA, Houghten R, Rhodes G, Vaughan JH;

XX WPI; 1992-226077/27.

XX Assay for anti-Epstein Barr virus nuclear antigen antibodies -
 PT using random copolymer polypeptide contg. specified aminoacid
 PT residues, for diagnosis of infectious mononucleosis and
 PT nasopharyngeal carcinoma

PS Claim 1; Page 54; 41pp; English.

XX The sequences given in AAR25050-7 are random copolymer peptides which
 CC can be used in diagnostic methods and systems relating to Epstein-
 CC Barr virus nuclear antigen (EBNA). These peptides are capable of
 CC inducing the production of antibodies which immunoreact with EBNA,
 CC and of immunoreacting with human antibodies induced by EBNA. The
 CC peptides can be fixed to a solid matrix, giving a solid
 CC immunoreactant in a diagnostic kit. The assay developed using
 CC these peptides has been found to be clinically reliable in
 CC detecting infectious mononucleosis (IM) caused by Epstein-Barr virus
 CC (EBV) as well as IM induced by cytomegalovirus, and also in

CC detecting nasopharyngeal carcinoma, another disease in which EBV
 CC has been implicated.
 CC (Updated on 25-MAR-2003 to correct PF field.)

XX Sequence 17 AA;

Query Match 1.9%; Score 7; DB 13; Length 17;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGG 240
 |||||
 Db 9 AGGGAGG 15

RESULT 70

AAAW79101

ID AAW79101 standard; peptide; 17 AA.

XX AAW79101;

DT 18-NOV-1998 (first entry)

DE Stabilising sequence of the invention.

XX Fusion protein; stabilising polypeptide; proteolytic degradation;
 KW resistance; half-life; autoimmune disease; inflammation; nitro drug;
 KW IkappaB regulator protein; inflammatory bowel disease; in vivo imaging;
 KW nitroreductase protein; enzyme therapy; prodrug therapy; protease;
 KW cancer; pathological condition.

XX Epstein-barr virus.

XX WO9822577-A1.

XX 28-MAY-1998.

XX 17-NOV-1997; 97WO-IB01508.

XX 25-JUN-1997; 97US-0048945.

PR 15-NOV-1996; 96US-0030986.

XX (MASU/) MASUCCI M G.

XX Masucci MG;

XX WPI; 1998-312463/27.

DR N-PSDB; AAV55829.

XX New fusion proteins resistant to proteolytic degradation -
 PT comprising a core protein with a stabilising polypeptide comprising
 PT a peptide sequence containing glycine repeats

XX Disclosure; Fig 2; 120pp; English.

XX This represents a stabilising sequence of the invention. The invention
 CC provides a method for increasing the resistance of a core protein to
 CC proteolytic degradation that comprises linking or inserting onto or into
 CC the core protein a stabilising polypeptide of formula [(Glya)(Glyb)y
 CC (Glyc)Z]n where Glya, Glyb, Glyc are 1-6 sequential Gly residues and
 CC x, y, z are Ala, Ser, Val, Ile, Leu, Met, Phe, Pro or Thr and n can be
 CC anything between 1-66. X, Y and Z need not be identical from n repeat to
 CC n repeat. Alternatively a nucleic acid encoding the stabilising
 CC polypeptide can be linked onto or inserted into a nucleic acid encoding
 CC a core protein. The fusion proteins of the invention are more resistant
 CC to degradation by proteases and, thus, have a longer half-life than the
 CC unfused core protein. The products can be used for treating autoimmune
 CC diseases, cancer and inflammation. In particular, the core protein may
 CC be an IkappaB regulator protein for the treatment of inflammatory bowel
 CC disease, or a nitroreductase protein which can activate nitro drugs in
 CC enzyme/prodrug therapy to treat cancer or other pathological conditions.
 CC The fusion proteins can also be used in diagnostic methods such as in
 CC vivo imaging.


```

XX SQ Sequence 17 AA;
Query Match 1.9%; Score 7; DB 19; Length 17;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGG 240
DB 11 AGGGAGG 17
|||||

RESULT 71
ABE76028
ID ABB76028 standard; Peptide; 19 AA.
XX AC ABB76028;
XX DT 12-JUL-2002 (first entry)
XX DE Streptococcus pneumoniae mucin-binding protein peptide fragment.
XX KW Mucin-binding protein; vaccine; diagnosis; gene therapy;
XX KW otitis media; rhinosinusitis; bacteraemia; meningitis; pneumonia;
XX KW lower respiratory tract infection; antimicrobial; antibiotic;
XX KW auditory; immunostimulant; vaccine.
XX OS Streptococcus pneumoniae.
XX PN WO200228351-A2.
XX PD 11-APR-2002.
XX PF 04-OCT-2001; 2001WO-US311269.
XX PR 04-OCT-2000; 2000US-237888P.
XX PR 07-FEB-2001; 2001US-267104P.
XX PA (AMHP ) AMERICAN HOME PROD CORP.
XX PA (UUNY ) UNIV NEW YORK STATE RES FOUND.
XX PI Green BA, Masi AW, Reddy MS;
XX WPI; 2002-383318/41.
XX DR Mucin binding proteins, useful in the induction of an immune response
XX PT to, and in the diagnosis of, pneumococcal infections -
XX PS Example; Page 69; 71pp; English.
XX CC The present sequence is an N-terminal peptide fragment of a 14 kDa
XX CC mucin-binding protein (see ABB76026) from the outer membrane
XX CC fraction of Streptococcus pneumoniae strain Cp1200. Elucidation
XX CC of this N-terminal sequence allowed the identification of the
XX CC corresponding open reading frame (see ABL54193) in the S.
XX CC pneumoniae genome. The invention provides amino acid and nucleic
XX CC acid sequences for isolated 12 kDa and 14 kDa mucin-binding proteins,
XX CC as well as expression vectors, transfected host cells, methods for
XX CC producing recombinant mucin-binding proteins, compositions comprising
XX CC the proteins, and antibodies to the proteins. Claimed compositions
XX CC comprising either a pneumococcal mucin-binding protein or an
XX CC expression vector encoding such a mucin-binding protein are used to
XX CC elicit protective immunity from a disease caused by S. pneumoniae,
XX CC especially otitis media, rhinosinusitis, bacteraemia, meningitis,
XX CC pneumonia or lower respiratory tract infection (all claimed).
XX SQ Sequence 19 AA;
Query Match 1.9%; Score 7; DB 23; Length 19;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 VKKNIEE 32

```

```

DB 9 VKKNIEE 15
|||||

RESULT 72
AAP60475
ID AAP60475 standard; protein; 20 AA.
XX AC AAP60475;
XX DT 25-MAR-2003 (updated)
XX DT 28-JUL-1991 (first entry)
XX DE Epstein-Barr virus nuclear antigen (EBNA) pentapeptide.
XX KW Epstein-Barr virus; nuclear antigen; passive immunization; diagnosis.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Region 2..6
XX FT /label= EBNA determinant
XX PN WO8601210-A.
XX PD 27-FEB-1986.
XX PF 02-AUG-1985; 85WO-US01484.
XX PR 08-AUG-1984; 84US-0638726.
XX PA (SCRI ) SCRIPPS CLINIC & RES FOUND.
XX PI Vaughan JH, Carbon DA, Rhodes G, Houghten R;
XX WPI; 1986-068962/10.
XX PT New synthetic copolymer polypeptide(s) - useful for preventing
XX PT and diagnosing diseases involving Epstein-Barr virus.
XX PS Claim 4(iii); Page 73; 84pp; English.
XX CC The peptide includes the EBNA pentapeptide which is an
XX CC immunogen for the production of an antibody in a host animal.
XX CC Epstein-Barr virus- and EBNA-associated diseases may be
XX CC diagnosed and prevented
XX CC See also AAP60473-P60483
XX CC (Updated on 25-MAR-2003 to correct PA field.)
XX SQ Sequence 20 AA;
Query Match 1.9%; Score 7; DB 7; Length 20;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGG 240
DB 3 AGGGAGG 9
|||||

RESULT 73
AAR03360
ID AAR03360 standard; protein; 20 AA.
XX AC AAR03360;
XX DT 25-MAR-2003 (updated)
XX DT 31-JUL-1990 (first entry)
XX DE Antigenic fragment of CMV-encoded polypeptide.
XX KW CMV; Epstein-Barr virus; EBV; infectious mononucleosis; IN;
XX KW nasopharyngeal carcinoma; NPC.

```

```

XX OS Cytomegalovirus.
XX PN WO9001495-A.
XX PD 22-FEB-1990.
XX PF 03-AUG-1989; 89WO-US03350.
XX PR 08-AUG-1988; 88US-0230250.
XX PA (SCRI ) SCRIPPS CLINIC & RE.
XX PI Rhodes G, Smith RS;
XX DR WPI; 1990-083477/11.
XX PT Synthetic peptide(s) corresp. to cytomegalovirus -
PT used to detect Epstein-Barr virus nuclear antigens and for
PT prodn. of antibodies.
XX PS Claim 3; Page 108; 135pp; English.
XX CC Peptide derivatives and antigenically related variants may be used as an
CC assay or to generate antibodies for passive immunisation against Epstein-
CC Barr virus encoded nuclear antigens, infectious mononucleosis and
CC nasopharyngeal carcinoma.
CC (Updated on 25-MAR-2003 to correct PF field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
XX SQ Sequence 20 AA;

Query Match 1.9%; Score 7; DB 11; Length 20;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 GGGAGGG 241
Db |||||
6 GGGAGGG 12

RESULT 74
AAR24430
ID AAR24430 standard; peptide; 20 AA.
XX AC AAR24430;
XX DT 21-NOV-1992 (first entry)
XX DE Sequence of synthetic Epstein-Barr virus (EBV)-encoded nuclear
XX antigen (EBNA) designated P62(B).
XX KW Herpes virus; Epstein-Barr virus; infectious mononucleosis;
XX Burkett's lymphoma; nasopharyngeal carcinoma; rheumatoid arthritis.
XX OS Synthetic.
XX PN US5116725-A.
XX PD 26-MAY-1992.
XX PF 04-NOV-1987; 87US-0117241.
XX PR 08-AUG-1984; 84US-0638726.
XX PR 24-MAR-1987; 87US-0029860.
XX PR 04-NOV-1987; 87US-0117241.
XX PA (SCRI ) SCRIPPS CLINIC & RES FOUND.
XX PI Carson DA, Geltosky JE, Houghten R, Rhodes G, Smith RS;
XX PI Vaughan JH;
XX DR WPI; 1992-199513/24.

```

```

XX PT Determining state of Epstein-Barr virus infection - by
XX PT determining ratio of IgG to IgM antibodies to synthetic peptides
XX PT in patient
XX PS Claim 1a1ii; Column 54; 40pp; English.
XX CC EBNA-1 is the earliest EBV-associated antigen that can be detected
XX CC after infection. EBNA has been detected in the nucleus of latently-
XX CC infected growth-transformed B lymphocytes. EBNA has also been
XX CC detected in the nuclei of African Burkitt tumor lymphoblasts and
XX CC anaplastic nasopharyngeal carcinoma cells. The size of the EBNA-1
XX CC protein is correlated with the variation of the length of the IR-3
XX CC region of EBV-DNA. The IR-3 region encodes a repeating glycine-
XX CC alanine sequence that has been characterised to be the major epitope
XX CC of the EBNA-1 protein. The peptides of the invention are synthetic,
XX CC random copolymeric polypeptides capable of inducing the production of
XX CC antibodies that immunoreact with EBNA and of immunoreacting with
XX CC human antibodies induced by EBNA. The sequences were chosen to
XX CC represent different areas from within and just outside the proposed
XX CC IR-3 region of the EBNA.
XX SQ Sequence 20 AA;

Query Match 1.9%; Score 7; DB 13; Length 20;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 240
Db |||||
3 AGGAGGG 9

RESULT 75
AAR25052
ID AAR25052 standard; Protein; 20 AA.
XX AC AAR25052;
XX DT 25-MAR-2003 (updated)
XX DT 09-DEC-1992 (first entry)
XX DE Synthetic random copolymer peptide P62.
XX KW EBNA; antibody; IM; EBV; cytomegalovirus.
XX OS Synthetic.
XX PN US5122448-A.
XX PD 16-JUN-1992.
XX PF 11-JAN-1990; 90US-0463505.
XX PR 08-AUG-1984; 84US-0638726.
XX PR 24-MAR-1987; 87US-0029860.
XX PR 04-NOV-1987; 87US-0117241.
XX PR 11-JAN-1990; 90US-0463505.
XX PA (SCRI ) SCRIPPS CLINIC & RES FOUND.
XX PI Carson DA, Houghten R, Rhodes G, Vaughan JH;
XX PI WPI; 1992-226077/27.
XX DR Assay for anti-Epstein Barr virus nuclear antigen antibodies -
XX PT using random copolymer polypeptide contg. specified aminoacid
XX PT residues, for diagnosis of infectious mononucleosis and
XX PT nasopharyngeal carcinoma
XX PS Claim 1; Page 54; 41pp; English.
XX DR The sequences given in AAR25050-7 are random copolymer peptides which

```

CC can be used in diagnostic methods and systems relating to Epstein-
CC Barr virus nuclear antigen (EBNA). These peptides are capable of
CC inducing the production of antibodies which immunoreact with EBNA,
CC and of immunoreacting with human antibodies induced by EBNA. The
CC peptides can be fixed to a solid matrix, giving a solid
CC immunoreactant in a diagnostic kit. The assay developed using
CC these peptides has been found to be clinically reliable in
CC detecting infectious mononucleosis (IM) caused by Epstein-Barr virus
CC (EBV) as well as IM induced by cytomegalovirus, and also in
CC detecting nasopharyngeal carcinoma, another disease in which EBV
CC has been implicated.
CC (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ Sequence 20 AA;

Query Match 1.9%; Score 7; DB 13; Length 20;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGG 240
| | | | |
Db 3 AGGGAGG 9

Search completed: December 17, 2003, 06:33:21
Job time : 37.7529 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 06:35:27 ; Search time 30.9444 Seconds
(without alignments)
2220.991 Million cell updates/sec

Title: US-09-805-550-4

Perfect score: 368

Sequence: 1 MKLTVKTLKGTFFIRVQPN.....CDRNEELAANYLLEHAGEED 368

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 696363 seqs, 186758610 residues

Word size : 0

Total number of hits satisfying chosen parameters: 696363

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	368	100.0	368	9	US-09-805-550-4
2	13	3.5	405	9	US-09-805-550-2
3	8	2.2	10	12	US-09-572-270A-6
4	8	2.2	10	12	US-09-572-270A-6
5	8	2.2	177	12	US-10-029-386-32842
6	8	2.2	306	12	US-10-267-430-13
7	8	2.2	325	12	US-10-267-430-4
8	8	2.2	331	12	US-10-194-885-10
9	8	2.2	332	12	US-10-091-841-9
10	8	2.2	401	15	US-10-295-403-148
11	8	2.2	744	12	US-10-063-685-67
12	8	2.2	744	15	US-10-184-644-255
13	8	2.2	744	15	US-10-184-634-255
14	8	2.2	755	12	US-10-140-472-153
15	8	2.2	755	12	US-10-141-761-153

16	8	2.2	755	12	US-10-142-885-153	Sequence 153, App
17	8	2.2	755	12	US-10-158-790-153	Sequence 153, App
18	8	2.2	755	12	US-10-137-871-153	Sequence 153, App
19	8	2.2	755	12	US-10-140-805-153	Sequence 153, App
20	8	2.2	755	12	US-10-140-864-153	Sequence 153, App
21	8	2.2	755	12	US-10-140-923-153	Sequence 153, App
22	8	2.2	755	12	US-10-141-756-153	Sequence 153, App
23	8	2.2	755	12	US-10-141-759-153	Sequence 153, App
24	8	2.2	755	15	US-10-123-155-153	Sequence 153, App
25	8	2.2	755	16	US-10-146-731-153	Sequence 153, App
26	8	2.2	828	12	US-10-174-677-34	Sequence 34, Appl
27	8	2.2	959	12	US-10-140-472-177	Sequence 177, App
28	8	2.2	959	12	US-10-141-761-177	Sequence 177, App
29	8	2.2	959	12	US-10-142-885-177	Sequence 177, App
30	8	2.2	959	12	US-10-158-790-177	Sequence 177, App
31	8	2.2	959	12	US-10-137-871-177	Sequence 177, App
32	8	2.2	959	12	US-10-140-805-177	Sequence 177, App
33	8	2.2	959	12	US-10-140-864-177	Sequence 177, App
34	8	2.2	959	12	US-10-140-923-177	Sequence 177, App
35	8	2.2	959	12	US-10-141-756-177	Sequence 177, App
36	8	2.2	959	12	US-10-141-759-177	Sequence 177, App
37	8	2.2	959	15	US-10-123-155-177	Sequence 177, App
38	8	2.2	959	16	US-10-146-731-177	Sequence 177, App
39	8	2.2	997	15	US-10-184-644-23	Sequence 23, Appl
40	8	2.2	1004	10	US-09-738-626-5676	Sequence 5676, Ap
41	8	2.2	1016	15	US-10-184-634-99	Sequence 99, Appl
42	8	2.2	1016	15	US-10-184-634-99	Sequence 99, Appl
43	8	2.2	1016	15	US-10-184-634-99	Sequence 99, Appl
44	8	2.2	1176	12	US-10-063-685-87	Sequence 87, Appl
45	8	2.2	1176	15	US-10-184-644-293	Sequence 293, App
46	8	2.2	1176	15	US-10-184-634-293	Sequence 293, App
47	8	2.2	1200	9	US-09-826-508-3	Sequence 3, Appl
48	8	2.2	1245	15	US-10-184-644-223	Sequence 223, App
49	8	2.2	1245	15	US-10-184-634-223	Sequence 223, App
50	8	2.2	1461	12	US-10-356-153-86	Sequence 86, Appl
51	8	2.2	1461	15	US-10-142-231-86	Sequence 86, Appl
52	8	2.2	1503	12	US-10-141-761-363	Sequence 363, App
53	8	2.2	1503	12	US-10-141-761-363	Sequence 363, App
54	8	2.2	1503	12	US-10-142-885-363	Sequence 363, App
55	8	2.2	1503	12	US-10-158-790-363	Sequence 363, App
56	8	2.2	1503	12	US-10-137-871-363	Sequence 363, App
57	8	2.2	1503	12	US-10-140-805-363	Sequence 363, App
58	8	2.2	1503	12	US-10-140-864-363	Sequence 363, App
59	8	2.2	1503	12	US-10-140-923-363	Sequence 363, App
60	8	2.2	1503	12	US-10-141-756-363	Sequence 363, App
61	8	2.2	1503	12	US-10-141-759-363	Sequence 363, App
62	8	2.2	1503	15	US-10-123-155-363	Sequence 363, App
63	8	2.2	1503	16	US-10-146-731-363	Sequence 363, App
64	8	2.2	1524	15	US-10-184-644-421	Sequence 421, App
65	8	2.2	1524	15	US-10-184-634-421	Sequence 421, App
66	8	2.2	1564	15	US-10-184-644-245	Sequence 245, App
67	8	2.2	1564	15	US-10-184-634-245	Sequence 245, App
68	8	2.2	1572	12	US-10-140-472-373	Sequence 373, App
69	8	2.2	1572	12	US-10-141-761-373	Sequence 373, App
70	8	2.2	1572	12	US-10-142-885-373	Sequence 373, App
71	8	2.2	1572	12	US-10-158-790-373	Sequence 373, App
72	8	2.2	1572	12	US-10-137-871-373	Sequence 373, App
73	8	2.2	1572	12	US-10-140-805-373	Sequence 373, App
74	8	2.2	1572	12	US-10-140-864-373	Sequence 373, App
75	8	2.2	1572	12	US-10-140-923-373	Sequence 373, App
76	8	2.2	1572	12	US-10-141-756-373	Sequence 373, App
77	8	2.2	1572	12	US-10-141-759-373	Sequence 373, App
78	8	2.2	1572	15	US-10-123-155-373	Sequence 373, App
79	8	2.2	1572	16	US-10-146-731-373	Sequence 373, App
80	8	2.2	1638	15	US-10-184-644-49	Sequence 49, Appl
81	8	2.2	1638	15	US-10-184-634-49	Sequence 49, Appl
82	8	2.2	1670	12	US-10-140-472-325	Sequence 325, App
83	8	2.2	1670	12	US-10-141-761-325	Sequence 325, App
84	8	2.2	1670	12	US-10-142-885-325	Sequence 325, App
85	8	2.2	1670	12	US-10-158-790-325	Sequence 325, App
86	8	2.2	1670	12	US-10-137-871-325	Sequence 325, App
87	8	2.2	1670	12	US-10-140-805-325	Sequence 325, App
88	8	2.2	1670	12	US-10-140-864-325	Sequence 325, App

; LENGTH: 10
; TYPE: PRT
; ORGANISM: Arabidopsis Thaliana
; OTHER INFORMATION: Sequence located in GRP7. at 138-147 and may interact with
US-09-572-270A-8

Query Match 2.2%; Score 8; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGGG 241
| | | | |
Db 1 AGGGAGGG 8

RESULT 5

US-10-029-386-32842
; Sequence 32842, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AECOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32842
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL035662.21
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: SWISSPROT HIT: P79955, EVALU2 2.00e-22
US-10-029-386-32842

Query Match 2.2%; Score 8; DB 12; Length 177;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGGG 241
| | | | |
Db 55 AGGGAGGG 62

RESULT 6

US-10-267-430-13
; Sequence 13, Application US/10267430
; Publication No. US20030175311A1
; GENERAL INFORMATION:
; APPLICANT: Binger, Mary-Helen
; APPLICANT: Chizzonite, Richard Anthony
; APPLICANT: Kramer, Richard Allen
; APPLICANT: Lomedico, Peter Thomas
; APPLICANT: McAndrew, Stephen J.
; APPLICANT: Altenburger, Werner
; TITLE OF INVENTION: Recombinant Coccidiosis Vaccines
; FILE REFERENCE: Recombinant Coccidiosis Vaccines
; CURRENT APPLICATION NUMBER: US/10/267,430
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US/08/744,793
; PRIOR FILING DATE: 1996-11-01
; PRIOR APPLICATION NUMBER: 07/812,349
; PRIOR FILING DATE: 1991-12-20
; PRIOR APPLICATION NUMBER: 07/202,721
; PRIOR FILING DATE: 1988-06-03

; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: AA seq. of the
; OTHER INFORMATION: nucleotide seq. of Fig. 27.
US-10-267-430-13

Query Match 2.2%; Score 8; DB 12; Length 306;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 233 NAGGGAGG 240
| | | | |
Db 206 NAGGGAGG 213

RESULT 7

US-10-267-430-4
; Sequence 4, Application US/10267430
; Publication No. US20030175311A1
; GENERAL INFORMATION:
; APPLICANT: Binger, Mary-Helen
; APPLICANT: Chizzonite, Richard Anthony
; APPLICANT: Kramer, Richard Allen
; APPLICANT: Lomedico, Peter Thomas
; APPLICANT: McAndrew, Stephen J.
; APPLICANT: Altenburger, Werner
; TITLE OF INVENTION: Recombinant Coccidiosis Vaccines
; FILE REFERENCE: Recombinant Coccidiosis Vaccines
; CURRENT APPLICATION NUMBER: US/10/267,430
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US/08/744,793
; PRIOR FILING DATE: 1996-11-01
; PRIOR APPLICATION NUMBER: 07/812,349
; PRIOR FILING DATE: 1991-12-20
; PRIOR APPLICATION NUMBER: 07/202,721
; PRIOR FILING DATE: 1988-06-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: AA seq of
; OTHER INFORMATION: nucleotide seq in Fig. 16.
US-10-267-430-4

Query Match 2.2%; Score 8; DB 12; Length 325;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 233 NAGGGAGG 240
| | | | |
Db 225 NAGGGAGG 232

RESULT 8

US-10-194-885-10
; Sequence 10, Application US/10194885
; Publication No. US20030135878A1
; GENERAL INFORMATION:
; APPLICANT: Wong, J. H.
; APPLICANT: Cho, Myeong-Je
; APPLICANT: Lemaux, Peggy G.
; APPLICANT: Buchanan, Bob
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH ELEVATED
; TITLE OF INVENTION: THIOREDUXIN LEVELS
; FILE REFERENCE: 41627200800

; CURRENT APPLICATION NUMBER: US/10/194,885
 ; CURRENT FILING DATE: 2002-07-12
 ; PRIOR APPLICATION NUMBER: 60/307,006
 ; PRIOR FILING DATE: 2001-07-19
 ; PRIOR APPLICATION NUMBER: 09/538,864
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: 60/126,736
 ; PRIOR FILING DATE: 1999-03-29
 ; NUMBER OF SEQ ID NOS: 55
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 10
 ; LENGTH: 331
 ; TYPE: PRT
 ; ORGANISM: Barley
 US-10-194-885-10

Query Match 2.2%; Score 8; DB 12; Length 331;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 GGAGGGPL 243
 |||||
 Db 221 GGAGGGPL 228

RESULT 9

US-10-091-841-9
 ; Sequence 9, Application US/10091841
 ; Publication No. US20030150010A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cho, Myeong-Je
 ; APPLICANT: Del Val, Greg
 ; APPLICANT: Caillaud, Maxine
 ; APPLICANT: Lemauz, Peggy G.
 ; APPLICANT: Buchanan, Bob B.
 ; TITLE OF INVENTION: Barley Gene for Thioresdoxin and
 ; TITLE OF INVENTION: NADP-Thioresdoxin Reductase
 ; FILE REFERENCE: 2001-0701.30
 ; CURRENT APPLICATION NUMBER: US/10/091,841
 ; CURRENT FILING DATE: 2002-03-05
 ; PRIOR APPLICATION NUMBER: 09/540,014
 ; PRIOR FILING DATE: 2000-03-31
 ; PRIOR APPLICATION NUMBER: US 60/127,198
 ; PRIOR FILING DATE: 1999-03-31
 ; PRIOR APPLICATION NUMBER: US 60/169,162
 ; PRIOR FILING DATE: 1998-12-06
 ; PRIOR APPLICATION NUMBER: US 60/177,740
 ; PRIOR FILING DATE: 2000-01-21
 ; PRIOR APPLICATION NUMBER: US 60/177,739
 ; PRIOR FILING DATE: 2000-01-21
 ; NUMBER OF SEQ ID NOS: 51
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9
 ; LENGTH: 332
 ; TYPE: PRT
 ; ORGANISM: Hordeum vulgare
 US-10-091-841-9

Query Match 2.2%; Score 8; DB 12; Length 332;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 GGAGGGPL 243
 |||||
 Db 221 GGAGGGPL 228

RESULT 10

US-10-295-403-148
 ; Sequence 148, Application US/10295403
 ; Publication No. US20030101481A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Heard, Jacqueline

; APPLICANT: Riechmann, Jose Luis
 ; APPLICANT: Adam, Luc
 ; APPLICANT: Broun, Pierre
 ; APPLICANT: Pineda, Onaira
 ; APPLICANT: Reuber, Lynne
 ; APPLICANT: Jiang, Cai-Zhong
 ; APPLICANT: Keddie, James
 ; APPLICANT: Zhang, James
 ; APPLICANT: Benito, Maria-Ines
 ; APPLICANT: Yu, Guo-Liang
 ; APPLICANT: Fromm, Mike
 ; TITLE OF INVENTION: PLANT GENE SEQUENCES I
 ; FILE REFERENCE: MBI-0003
 ; CURRENT APPLICATION NUMBER: US/10/295,403
 ; CURRENT FILING DATE: 2002-11-15
 ; PRIOR APPLICATION NUMBER: US/09/394,519
 ; PRIOR FILING DATE: 1999-09-13
 ; PRIOR APPLICATION NUMBER: 60/101,349
 ; PRIOR FILING DATE: 1998-09-22
 ; PRIOR APPLICATION NUMBER: 60/103,312
 ; PRIOR FILING DATE: 1998-10-06
 ; PRIOR APPLICATION NUMBER: 60/108,734
 ; PRIOR FILING DATE: 1998-11-17
 ; PRIOR APPLICATION NUMBER: 60/113,409
 ; PRIOR FILING DATE: 1998-12-22
 ; NUMBER OF SEQ ID NOS: 170
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 148
 ; LENGTH: 401
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; OTHER INFORMATION: G802
 US-10-295-403-148

Query Match 2.2%; Score 8; DB 15; Length 401;
 Best Local Similarity 100.0%; Pred. No. 54;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGGG 241
 |||||
 Db 294 AGGAGGG 301

RESULT 11

US-10-063-685-67
 ; Sequence 67, Application US/10063685
 ; Publication No. US20030180909A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Geritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; TITLE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3230R1C1
 ; CURRENT APPLICATION NUMBER: US/10/063,685
 ; CURRENT FILING DATE: 2002-05-08
 ; PRIOR APPLICATION removed - See Palm or File Wrapper
 ; NUMBER OF SEQ ID NOS: 170
 ; SEQ ID NO 67
 ; LENGTH: 744
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-063-685-67

Query Match 2.2%; Score 8; DB 12; Length 744;
 Best Local Similarity 100.0%; Pred. No. 94;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
 Db 17 AGGGAGGG 24

RESULT 12

US-10-184-644-255
 ; Sequence 255, Application US/10184644
 ; Publication No. US20030044930A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3430R1C227
 ; CURRENT APPLICATION NUMBER: US/10/184,644
 ; CURRENT FILING DATE: 2002-06-28
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 255
 ; LENGTH: 744
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-184-644-255

Query Match 2.2%; Score 8; DB 15; Length 744;
 Best Local Similarity 100.0%; Pred. No. 94;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
 Db 17 AGGGAGGG 24

RESULT 13

US-10-184-634-255
 ; Sequence 255, Application US/10184634
 ; Publication No. US20030068684A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3430R1C217
 ; CURRENT APPLICATION NUMBER: US/10/184,634
 ; CURRENT FILING DATE: 2002-06-28
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 255
 ; LENGTH: 744
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-184-634-255

Query Match 2.2%; Score 8; DB 15; Length 744;
 Best Local Similarity 100.0%; Pred. No. 94;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
 Db 17 AGGGAGGG 24

RESULT 14

US-10-140-472-153
 ; Sequence 153, Application US/10140472
 ; Publication No. US20030138888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3330R1C168
 ; CURRENT APPLICATION NUMBER: US/10/140,472
 ; CURRENT FILING DATE: 2002-05-06
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 153
 ; LENGTH: 755
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: 67
 ; OTHER INFORMATION: unknown base
 US-10-140-472-153

Query Match 2.2%; Score 8; DB 12; Length 755;
 Best Local Similarity 100.0%; Pred. No. 95;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
 Db 634 AGGGAGGG 641

RESULT 15

US-10-141-761-153
 ; Sequence 153, Application US/10141761
 ; Publication No. US20030148432A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood,Steven
 ; APPLICANT: Smith,Victoria
 ; APPLICANT: Stewart,Timothy A.
 ; APPLICANT: Tumas,Daniel
 ; APPLICANT: Watanabe,Colin K
 ; APPLICANT: Wood,William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3330R1C198
 ; CURRENT APPLICATION NUMBER: US/10/141,761
 ; CURRENT FILING DATE: 2002-05-08
 ; Prior Application removed - See Palm or File Wrapper
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 153
 ; LENGTH: 755
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: 67
 ; OTHER INFORMATION: unknown base
 US-10-141-761-153

Query Match 2.2%; Score 8; DB 12; Length 755;
 Best Local Similarity 100.0%; Pred. No. 95;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
 Db 634 AGGGAGGG 641

RESULT 16

US-10-142-885-153
 ; Sequence 153, Application US/10142885
 ; Publication No. US20030157604A1
 ; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3330R1C248

; CURRENT APPLICATION NUMBER: US/10/142,885
 ; CURRENT FILING DATE: 2002-05-10

; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 153
 ; LENGTH: 755

; TYPE: DNA
 ; ORGANISM: Homo Sapien
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: 67
 ; OTHER INFORMATION: unknown base

US-10-142-885-153

Query Match 2.2%; Score 8; DB 12; Length 755;

Best Local Similarity 100.0%; Pred. No. 95;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
 Db 634 AGGGAGGG 641

RESULT 17

US-10-158-790-153
 ; Sequence 153, Application US/10158790

; Publication No. US20030180879A1
 ; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3330R1C448

; CURRENT APPLICATION NUMBER: US/10/158,790
 ; CURRENT FILING DATE: 2002-05-30

; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 153
 ; LENGTH: 755

; TYPE: DNA
 ; ORGANISM: Homo Sapien
 ; FEATURE:
 ; NAME/KEY: unsure

; LOCATION: 67
 ; OTHER INFORMATION: unknown base

US-10-158-790-153

Query Match 2.2%; Score 8; DB 12; Length 755;
 Best Local Similarity 100.0%; Pred. No. 95;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
 Db 634 AGGGAGGG 641

RESULT 18

US-10-137-871-153

; Sequence 153, Application US/10137871
 ; Publication No. US20030207350A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria

```
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: F3330R1C153
; CURRENT APPLICATION NUMBER: US/10/137,871
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 153
; LENGTH: 755
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 67
; OTHER INFORMATION: unknown base
; US-10-137-871-153

Query Match      2.2%; Score 8; DB 12; Length 755;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      234 AGGGAGGG 241
Db      634 AGGGAGGG 641
|||||

RESULT 19
US-10-140-805-153
; Sequence 153, Application US/10140805
; Publication No. US20030207417A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: F3330R1C176
; CURRENT APPLICATION NUMBER: US/10/140,805
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 153
; LENGTH: 755
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 67
; OTHER INFORMATION: unknown base
; US-10-140-805-153

Query Match      2.2%; Score 8; DB 12; Length 755;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      234 AGGGAGGG 241
Db      634 AGGGAGGG 641
|||||

RESULT 20
US-10-140-864-153
; Sequence 153, Application US/10140864
; Publication No. US20030207419A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: F3330R1C184
; CURRENT APPLICATION NUMBER: US/10/140,864
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 153
; LENGTH: 755
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 67
; OTHER INFORMATION: unknown base
; US-10-140-864-153

Query Match      2.2%; Score 8; DB 12; Length 755;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      234 AGGGAGGG 241
Db      634 AGGGAGGG 641
|||||

RESULT 21
US-10-140-923-153
; Sequence 153, Application US/10140923
; Publication No. US20030207355A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
```

```
/ APPLICANT: Watanabe, Colin K
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3330RIC188
/ CURRENT APPLICATION NUMBER: US/10/140,923
/ CURRENT FILING DATE: 2002-05-07
/ Prior Application removed - See Palm or File Wrapper
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 153
/ LENGTH: 755
/ TYPE: DNA
/ ORGANISM: Homo Sapien
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 67
/ OTHER INFORMATION: unknown base
US-10-140-923-153

Query Match      2.2%; Score 8; DB 12; Length 755;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      234 AGGGAGGG 241
Db      634 AGGGAGGG 641
|||||

RESULT 22
US-10-141-756-153
/ Sequence 153, Application US/10141756
/ Publication No. US20030207359A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Beresini, Maureen
/ APPLICANT: DeForge, Laura
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3330RIC197
/ CURRENT APPLICATION NUMBER: US/10/141,759
/ CURRENT FILING DATE: 2002-05-08
/ Prior Application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 153
/ LENGTH: 755
/ TYPE: DNA
/ ORGANISM: Homo Sapien
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 67
/ OTHER INFORMATION: unknown base
US-10-141-759-153

Query Match      2.2%; Score 8; DB 12; Length 755;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      234 AGGGAGGG 241
Db      634 AGGGAGGG 641
|||||

RESULT 24
US-10-123-155-153
/ Sequence 153, Application US/10123155
/ Publication No. US20030068794A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Beresini, Maureen
/ APPLICANT: DeForge, Laura
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K
/ APPLICANT: Wood, William
```

```
/ APPLICANT: Watanabe, Colin K
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3330RIC188
/ CURRENT APPLICATION NUMBER: US/10/140,923
/ CURRENT FILING DATE: 2002-05-07
/ Prior Application removed - See Palm or File Wrapper
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 153
/ LENGTH: 755
/ TYPE: DNA
/ ORGANISM: Homo Sapien
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 67
/ OTHER INFORMATION: unknown base
US-10-140-923-153

Query Match      2.2%; Score 8; DB 12; Length 755;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      234 AGGGAGGG 241
Db      634 AGGGAGGG 641
|||||

RESULT 22
US-10-141-756-153
/ Sequence 153, Application US/10141756
/ Publication No. US20030207359A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Beresini, Maureen
/ APPLICANT: DeForge, Laura
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3330RIC200
/ CURRENT APPLICATION NUMBER: US/10/141,756
/ CURRENT FILING DATE: 2002-05-08
/ Prior Application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 153
/ LENGTH: 755
/ TYPE: DNA
/ ORGANISM: Homo Sapien
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 67
/ OTHER INFORMATION: unknown base
US-10-141-756-153

Query Match      2.2%; Score 8; DB 12; Length 755;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      234 AGGGAGGG 241
```

; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: F330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 153
; LENGTH: 755
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 67
; OTHER INFORMATION: unknown base
US-10-123-155-153

Query Match 2.2%; Score 8; DB 15; Length 755;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 241
| | | | |
Db 634 AGGAGGG 641

RESULT 25
US-10-146-731-153
; Sequence 153, Application US/10146731
; Publication No. US20030129692A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: F330R1C323
; CURRENT APPLICATION NUMBER: US/10/146,731
; CURRENT FILING DATE: 2002-05-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 153
; LENGTH: 755
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 67
; OTHER INFORMATION: unknown base
US-10-146-731-153

Query Match 2.2%; Score 8; DB 16; Length 755;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 241
| | | | |
Db 634 AGGAGGG 641

RESULT 26

US-10-174-677-34
; Sequence 34, Application US/10174677
; Publication No. US20030190704A1
; GENERAL INFORMATION:

; APPLICANT: Xie, Ting
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ANCHORING STEM CELLS IN A MICROENV
; FILE REFERENCE: 40716(IP-012)
; CURRENT APPLICATION NUMBER: US/10/174,677
; CURRENT FILING DATE: 2002-06-19
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 828
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-174-677-34

Query Match 2.2%; Score 8; DB 12; Length 828;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 241
| | | | |
Db 706 AGGAGGG 713

RESULT 27

US-10-140-472-177
; Sequence 177, Application US/10140472
; Publication No. US2003013888A1
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: F330R1C168
; CURRENT APPLICATION NUMBER: US/10/140,472
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 177
; LENGTH: 959
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-472-177

Query Match 2.2%; Score 8; DB 12; Length 959;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 241
| | | | |
Db 676 AGGAGGG 683

RESULT 28

US-10-141-761-177
; Sequence 177, Application US/10141761
; Publication No. US20030148432A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C198
; CURRENT APPLICATION NUMBER: US/10/141,761
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 177
; LENGTH: 959
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-141-761-177

Query Match 2.2%; Score 8; DB 12; Length 959;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
Db 676 AGGGAGGG 683

RESULT 29

US-10-142-885-177
; Sequence 177, Application US/10142885
; Publication No. US20030157604A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C248
; CURRENT APPLICATION NUMBER: US/10/142,885
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 177
; LENGTH: 959
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-142-885-177

Query Match 2.2%; Score 8; DB 12; Length 959;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
Db 676 AGGGAGGG 683

RESULT 30

US-10-158-790-177
; Sequence 177, Application US/10158790
; Publication No. US20030180879A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C448
; CURRENT APPLICATION NUMBER: US/10/158,790
; CURRENT FILING DATE: 2002-05-30
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 177
; LENGTH: 959
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-158-790-177

Query Match 2.2%; Score 8; DB 12; Length 959;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
Db 676 AGGGAGGG 683

RESULT 31

US-10-137-871-177
; Sequence 177, Application US/10137871
; Publication No. US20030207350A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.

```
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C153
; CURRENT APPLICATION NUMBER: US/10/137,871
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 177
; LENGTH: 959
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-137-871-177

Query Match      2.2%; Score 8; DB 12; Length 959;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      234 AGGGAGGG 241
Db      676 AGGGAGGG 683
|||||

RESULT 32
US-10-140-805-177
; Sequence 177, Application US/10140805
; Publication No. US20030207417A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C176
; CURRENT APPLICATION NUMBER: US/10/140,805
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 177
; LENGTH: 959
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-805-177

Query Match      2.2%; Score 8; DB 12; Length 959;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      234 AGGGAGGG 241
Db      676 AGGGAGGG 683
|||||
```

```
RESULT 33
US-10-140-864-177
; Sequence 177, Application US/10140864
; Publication No. US20030207419A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C184
; CURRENT APPLICATION NUMBER: US/10/140,864
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 177
; LENGTH: 959
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-864-177

Query Match      2.2%; Score 8; DB 12; Length 959;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      234 AGGGAGGG 241
Db      676 AGGGAGGG 683
|||||

RESULT 34
US-10-140-923-177
; Sequence 177, Application US/10140923
; Publication No. US20030207355A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C188
; CURRENT APPLICATION NUMBER: US/10/140,923
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See Palm or File Wrapper
```

; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 177
 ; LENGTH: 959
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-140-923-177

Query Match 2.2%; Score 8; DB 12; Length 959;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGGG 241
 |||||
 Db 676 AGGGAGGG 683

RESULT 35

US-10-141-756-177
 ; Sequence 177, Application US/10141756
 ; Publication No. US20030207359A1
 ; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3330R1C200
 ; CURRENT APPLICATION NUMBER: US/10/141,756
 ; CURRENT FILING DATE: 2002-05-08
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 177
 ; LENGTH: 959
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-141-756-177

Query Match 2.2%; Score 8; DB 12; Length 959;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGGG 241
 |||||
 Db 676 AGGGAGGG 683

RESULT 36

US-10-141-759-177
 ; Sequence 177, Application US/10141759
 ; Publication No. US20030207361A1
 ; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3330R1C197
 ; CURRENT APPLICATION NUMBER: US/10/141,759
 ; CURRENT FILING DATE: 2002-05-08
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 177
 ; LENGTH: 959
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-141-759-177

Query Match 2.2%; Score 8; DB 12; Length 959;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 234 AGGGAGGG 241
 |||||
 Db 676 AGGGAGGG 683

RESULT 37

US-10-123-155-177
 ; Sequence 177, Application US/10123155
 ; Publication No. US20030068794A1
 ; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3330R1C30
 ; CURRENT APPLICATION NUMBER: US/10/123,155
 ; CURRENT FILING DATE: 2002-04-15
 ; Prior Application removed - See Palm or File Wrapper
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 177
 ; LENGTH: 959
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-123-155-177

Query Match 2.2%; Score 8; DB 15; Length 959;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGGG 241
 |||||
 Db 676 AGGGAGGG 683

; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 5676
 ; LENGTH: 1004
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-5676

Query Match 2.2%; Score 8; DB 10; Length 1004;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 241
 |||||
 Db 338 AGGAGGG 345

RESULT 42

US-10-184-644-99
 ; Sequence 99, Application US/10184644
 ; Publication No. US20030044930A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; TITLE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3430R1C227
 ; CURRENT APPLICATION NUMBER: US/10/184,644
 ; CURRENT FILING DATE: 2002-06-28
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 99
 ; LENGTH: 1016
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-184-644-99

Query Match 2.2%; Score 8; DB 15; Length 1016;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 241
 |||||
 Db 667 AGGAGGG 674

RESULT 43

US-10-184-634-99
 ; Sequence 99, Application US/10184634
 ; Publication No. US20030068684A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3430R1C217
 ; CURRENT APPLICATION NUMBER: US/10/184,634
 ; CURRENT FILING DATE: 2002-06-28
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 99
 ; LENGTH: 1016
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-184-634-99

Query Match 2.2%; Score 8; DB 15; Length 1016;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 241
 |||||
 Db 667 AGGAGGG 674

RESULT 44

US-10-063-685-87
 ; Sequence 87, Application US/10063685
 ; Publication No. US20030180909A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; TITLE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3230R1C1
 ; CURRENT APPLICATION NUMBER: US/10/063,685
 ; CURRENT FILING DATE: 2002-05-08
 ; Prior Application removed - See Palm or File Wrapper
 ; NUMBER OF SEQ ID NOS: 170
 ; SEQ ID NO 87
 ; LENGTH: 1176
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-063-685-87

Query Match 2.2%; Score 8; DB 12; Length 1176;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 241
 |||||
 Db 38 AGGAGGG 45

RESULT 45

US-10-184-644-293
 ; Sequence 293, Application US/10184644
 ; Publication No. US20030044930A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin

```
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C227
; CURRENT APPLICATION NUMBER: US/10/184,644
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 293
; LENGTH: 1176
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-644-293

Query Match      2.2%; Score 8; DB 15; Length 1176;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      234 AGGGAGGG 241
Db      38 AGGGAGGG 45

RESULT 46
US-10-184-634-293
; Sequence 293, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C217
; CURRENT APPLICATION NUMBER: US/10/184,634
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 293
; LENGTH: 1176
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-634-293

Query Match      2.2%; Score 8; DB 15; Length 1176;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      234 AGGGAGGG 241
Db      38 AGGGAGGG 45

RESULT 47
US-09-826-508-3
; Sequence 3, Application US/09826508
; Patent No. US20010025099A1
; GENERAL INFORMATION:
; APPLICANT: Nabil Elshourbagy
; APPLICANT: Lisa Vawter
; TITLE OF INVENTION: G Protein-Coupled Receptor Polypeptides
; FILE REFERENCE: GP-70744USB
; CURRENT APPLICATION NUMBER: US/09/826,508
; CURRENT FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 40
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1200
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-826-508-3

Query Match      2.2%; Score 8; DB 9; Length 1200;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      234 AGGGAGGG 241
Db      1087 AGGGAGGG 1094

RESULT 48
US-10-184-644-223
; Sequence 223, Application US/10184644
; Publication No. US2003004930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C227
; CURRENT APPLICATION NUMBER: US/10/184,644
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 223
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-644-223

Query Match      2.2%; Score 8; DB 15; Length 1245;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      234 AGGGAGGG 241
Db      38 AGGGAGGG 45

RESULT 49
US-10-184-634-223
; Sequence 223, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C217
```

; CURRENT APPLICATION NUMBER: US/10/184,634
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 223
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-634-223

Query Match 2.2%; Score 8; DB 15; Length 1245;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 241
Db 38 AGGAGGG 45
|||||

RESULT 50
US-10-356-153-86
; Sequence 86, Application US/10356153
; Publication No. US20030166176A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: CYTOCHROME P450 OXYGENASES AND THEIR USES
; FILE REFERENCE: 62773
; CURRENT APPLICATION NUMBER: US/10/356,153
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: US/10/142,231
; PRIOR FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: 60/165,250
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
; LENGTH: 1461
; TYPE: PRT
; ORGANISM: Taxus cuspidata
US-10-356-153-86

Query Match 2.2%; Score 8; DB 12; Length 1461;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 241
Db 817 AGGAGGG 824
|||||

RESULT 51
US-10-142-231-86
; Sequence 86, Application US/10142231
; Publication No. US2003007796A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: CYTOCHROME P450 OXYGENASES AND THEIR USES
; FILE REFERENCE: 62773
; CURRENT APPLICATION NUMBER: US/10/142,231
; CURRENT FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: 60/165,250
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
; LENGTH: 1461
; TYPE: PRT
; ORGANISM: Taxus cuspidata
US-10-142-231-86

Query Match 2.2%; Score 8; DB 15; Length 1461;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 241
Db 817 AGGAGGG 824
|||||

RESULT 52
US-10-140-472-363
; Sequence 363, Application US/10140472
; Publication No. US2003013888A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C168
; CURRENT APPLICATION NUMBER: US/10/140,472
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 363
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-472-363

Query Match 2.2%; Score 8; DB 12; Length 1503;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 241
Db 1203 AGGAGGG 1210
|||||

RESULT 53
US-10-141-761-363
; Sequence 363, Application US/10141761
; Publication No. US20030148432A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3330R1C198
 ; CURRENT APPLICATION NUMBER: US/10/141,761
 ; CURRENT FILING DATE: 2002-05-08
 ; Prior Application removed - See Palm or File Wrapper
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 363
 ; LENGTH: 1503
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-141-761-363

Query Match 2.2%; Score 8; DB 12; Length 1503;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
 |||||
 Db 1203 AGGGAGGG 1210

RESULT 54
 US-10-142-885-363
 ; Sequence 363, Application US/10142885
 ; Publication No. US20030157604A1
 ; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3330R1C248
 ; CURRENT APPLICATION NUMBER: US/10/142,885
 ; CURRENT FILING DATE: 2002-05-10
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 363
 ; LENGTH: 1503
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-142-885-363

Query Match 2.2%; Score 8; DB 12; Length 1503;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
 |||||
 Db 1203 AGGGAGGG 1210

RESULT 55
 US-10-158-790-363
 ; Sequence 363, Application US/10158790
 ; Publication No. US20030180879A1
 ; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3330R1C448
 ; CURRENT APPLICATION NUMBER: US/10/158,790
 ; CURRENT FILING DATE: 2002-05-30
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 363
 ; LENGTH: 1503
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-158-790-363

Query Match 2.2%; Score 8; DB 12; Length 1503;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241
 |||||
 Db 1203 AGGGAGGG 1210

RESULT 56
 US-10-137-871-363
 ; Sequence 363, Application US/10137871
 ; Publication No. US20030207350A1
 ; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3330R1C153
 ; CURRENT APPLICATION NUMBER: US/10/137,871
 ; CURRENT FILING DATE: 2002-05-03
 ; Prior Application removed - See Palm or File Wrapper
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 363
 ; LENGTH: 1503
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-137-871-363

Query Match 2.2%; Score 8; DB 12; Length 1503;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGGG 241
|||||||
Db 1203 AGGAGGG 1210

RESULT 57

US-10-140-805-363
; Sequence 363, Application US/10140805
; Publication No. US20030207417A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C176
; CURRENT APPLICATION NUMBER: US/10/140,805
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 363
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-805-363

Query Match 2.2%; Score 8; DB 12; Length 1503;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGGG 241
|||||||
Db 1203 AGGAGGG 1210

RESULT 58

US-10-140-864-363
; Sequence 363, Application US/10140864
; Publication No. US20030207419A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C184
; CURRENT APPLICATION NUMBER: US/10/140,864
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 363
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-864-363

Query Match 2.2%; Score 8; DB 12; Length 1503;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGGG 241
|||||||
Db 1203 AGGAGGG 1210

RESULT 59

US-10-140-923-363
; Sequence 363, Application US/10140923
; Publication No. US20030207355A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C188
; CURRENT APPLICATION NUMBER: US/10/140,923
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 363
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-923-363

Query Match 2.2%; Score 8; DB 12; Length 1503;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGGG 241
|||||||
Db 1203 AGGAGGG 1210

RESULT 60

US-10-141-756-363
; Sequence 363, Application US/10141756
; Publication No. US20030207359A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen

```

; APPLICANT: DeForge,Laura
; APPLICANT: Desnoyers,Luc
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C200
; CURRENT APPLICATION NUMBER: US/10/141,756
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 363
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-141-756-363

Query Match      2.2%; Score 8; DB 12; Length 1503;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      234 AGGGAGGG 241
      |||||
Db      1203 AGGGAGGG 1210

RESULT 61
US-10-141-759-363
; Sequence 363, Application US/10141759
; Publication No. US20030207361A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C197
; CURRENT APPLICATION NUMBER: US/10/141,759
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 363
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-141-759-363

Query Match      2.2%; Score 8; DB 12; Length 1503;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      234 AGGGAGGG 241
      |||||
Db      1203 AGGGAGGG 1210

US-10-123-155-363
; Sequence 363, Application US/10123155
; Publication No. US20030068794A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 363
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-363

Query Match      2.2%; Score 8; DB 15; Length 1503;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      234 AGGGAGGG 241
      |||||
Db      1203 AGGGAGGG 1210

RESULT 63
US-10-146-731-363
; Sequence 363, Application US/10146731
; Publication No. US20030129692A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William

```

```

Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      234 AGGGAGGG 241
      |||||
Db      1203 AGGGAGGG 1210

RESULT 62
US-10-123-155-363
; Sequence 363, Application US/10123155
; Publication No. US20030068794A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 363
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-363

Query Match      2.2%; Score 8; DB 15; Length 1503;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      234 AGGGAGGG 241
      |||||
Db      1203 AGGGAGGG 1210

RESULT 63
US-10-146-731-363
; Sequence 363, Application US/10146731
; Publication No. US20030129692A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William

```

```
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C323
; CURRENT APPLICATION NUMBER: US/10/146,731
; CURRENT FILING DATE: 2002-05-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 363
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-146-731-363

Query Match      2.2%; Score 8; DB 16; Length 1503;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      234 AGGGAGGG 241
Db      1203 AGGGAGGG 1210

RESULT 64
US-10-184-644-421
; Sequence 421, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C227
; CURRENT APPLICATION NUMBER: US/10/184,644
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 421
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-644-421

Query Match      2.2%; Score 8; DB 15; Length 1524;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      234 AGGGAGGG 241
Db      923 AGGGAGGG 930

RESULT 65
US-10-184-634-421
; Sequence 421, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
```

```
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C217
; CURRENT APPLICATION NUMBER: US/10/184,634
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 421
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-634-421

Query Match      2.2%; Score 8; DB 15; Length 1524;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      234 AGGGAGGG 241
Db      923 AGGGAGGG 930

RESULT 66
US-10-184-644-245
; Sequence 245, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C227
; CURRENT APPLICATION NUMBER: US/10/184,644
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 245
; LENGTH: 1564
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-644-245

Query Match      2.2%; Score 8; DB 15; Length 1564;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      234 AGGGAGGG 241
Db      903 AGGGAGGG 910

RESULT 67
US-10-184-634-245
; Sequence 245, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
```

```

; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC217
; CURRENT APPLICATION NUMBER: US/10/184,634
; PRIOR FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 245
; LENGTH: 1564
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-634-245

Query Match      2.2%; Score 8; DB 15; Length 1564;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      234 AGGGAGGG 241
Db      903 AGGGAGGG 910

RESULT 68
US-10-140-472-373
; Sequence 373, Application US/10140472
; Publication No. US2003013888A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC168
; CURRENT APPLICATION NUMBER: US/10/140,472
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 373
; LENGTH: 1572
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-472-373

Query Match      2.2%; Score 8; DB 12; Length 1572;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      234 AGGGAGGG 241
Db      709 AGGGAGGG 716

US-10-140-472-373
; Sequence 373, Application US/10140472
; Publication No. US2003013888A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC168
; CURRENT APPLICATION NUMBER: US/10/140,472
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 373
; LENGTH: 1572
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-472-373

Query Match      2.2%; Score 8; DB 12; Length 1572;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      234 AGGGAGGG 241
Db      709 AGGGAGGG 716

```

```

RESULT 69
US-10-141-761-373
; Sequence 373, Application US/10141761
; Publication No. US20030148432A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC198
; CURRENT APPLICATION NUMBER: US/10/141,761
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 373
; LENGTH: 1572
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-141-761-373

Query Match      2.2%; Score 8; DB 12; Length 1572;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      234 AGGGAGGG 241
Db      709 AGGGAGGG 716

RESULT 70
US-10-142-885-373
; Sequence 373, Application US/10142885
; Publication No. US20030157604A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC248
; CURRENT APPLICATION NUMBER: US/10/142,885
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550

```


; SEQ ID NO 373
; LENGTH: 1572
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-142-885-373

Query Match 2.2%; Score 8; DB 12; Length 1572;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 241
Db 709 AGGAGGG 716
|||||

RESULT 71

US-10-158-790-373
; Sequence 373, Application US/10158790
; Publication No. US20030180879A1
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C448
; CURRENT APPLICATION NUMBER: US/10/158,790
; CURRENT FILING DATE: 2002-05-30
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 373
; LENGTH: 1572
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-158-790-373

Query Match 2.2%; Score 8; DB 12; Length 1572;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 241
Db 709 AGGAGGG 716
|||||

RESULT 72

US-10-137-871-373
; Sequence 373, Application US/10137871
; Publication No. US20030207350A1
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C153
; CURRENT APPLICATION NUMBER: US/10/137,871
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 373
; LENGTH: 1572
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-137-871-373

Query Match 2.2%; Score 8; DB 12; Length 1572;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 241
Db 709 AGGAGGG 716
|||||

RESULT 73

US-10-140-805-373
; Sequence 373, Application US/10140805
; Publication No. US20030207417A1
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C176
; CURRENT APPLICATION NUMBER: US/10/140,805
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 373
; LENGTH: 1572
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-805-373

Query Match 2.2%; Score 8; DB 12; Length 1572;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGGG 241
Db 709 AGGAGGG 716
|||||

RESULT 74

US-10-140-864-373
; Sequence 373, Application US/10140864
; Publication No. US20030207419A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3330R1C184

; CURRENT APPLICATION NUMBER: US/10/140,864
; CURRENT FILING DATE: 2002-05-07

; Prior Application removed - See Palm or File Wrapper

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 373

; LENGTH: 1572

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-140-864-373

Query Match

2.2%; Score 8; DB 12; Length 1572;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241

|||||

Db 709 AGGGAGGG 716

RESULT 75

US-10-140-923-373

; Sequence 373, Application US/10140923

; Publication No. US20030207355A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3330R1C188

; CURRENT APPLICATION NUMBER: US/10/140,923
; CURRENT FILING DATE: 2002-05-07

; Prior Application removed - See Palm or File Wrapper

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 373

; LENGTH: 1572

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-140-923-373

Query Match

2.2%; Score 8; DB 12; Length 1572;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGGG 241

|||||

Db 709 AGGGAGGG 716

Search completed: December 17, 2003, 06:43:01

Job time : 32.9444 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 06:32:02 ; Search time 12.8538 Seconds
(without alignments)
1211.344 Million cell updates/sec

Title: US-09-805-550-4

Perfect score: 368

Sequence: 1 MKLTVKTLKGTHFEIRVQPN.....CORNEELAANYLLEHAGED 368

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A-COMB.pep:*
- 2: /cgn2_6/ptodata/1/iaa/5B-COMB.pep:*
- 3: /cgn2_6/ptodata/1/iaa/6A-COMB.pep:*
- 4: /cgn2_6/ptodata/1/iaa/6B-COMB.pep:*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS-COMB.pep:*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	368	100.0	368	3	US-09-413-574-4
2	13	3.5	405	3	US-09-413-574-2
3	8	2.2	332	4	US-09-540-014-9
4	8	2.2	371	4	US-09-252-991A-25766
5	8	2.2	379	4	US-09-252-991A-28428
6	8	2.2	738	3	US-08-864-038A-3
7	8	2.2	1417	4	US-08-900-230-3
8	7	1.9	20	3	US-08-961-308B-5
9	7	1.9	57	3	US-08-506-553C-1
10	7	1.9	61	4	US-09-149-476-615
11	7	1.9	81	4	US-09-249-542-9
12	7	1.9	82	4	US-09-249-542-8
13	7	1.9	96	3	US-09-036-113-2
14	7	1.9	96	4	US-09-556-166-2
15	7	1.9	96	4	US-09-252-991A-29877
16	7	1.9	111	4	US-09-249-542-16
17	7	1.9	111	4	US-09-249-542-17
18	7	1.9	111	4	US-09-249-542-18
19	7	1.9	112	4	US-09-249-542-7
20	7	1.9	112	4	US-09-249-542-10
21	7	1.9	112	4	US-09-249-542-11
22	7	1.9	112	4	US-09-249-542-14
23	7	1.9	112	4	US-09-249-542-15
24	7	1.9	112	4	US-09-249-542-19
25	7	1.9	112	4	US-09-249-542-24
26	7	1.9	128	4	US-09-996-243-420
27	7	1.9	160	4	US-09-328-352-6723

28	7	1.9	161	2	US-08-403-852D-30	Sequence 30, Appl
29	7	1.9	161	3	US-08-510-646B-31	Sequence 31, Appl
30	7	1.9	161	3	US-09-231-818-30	Sequence 30, Appl
31	7	1.9	185	4	US-09-562-737-112	Sequence 112, Appl
32	7	1.9	208	2	US-08-606-143-40	Sequence 40, Appl
33	7	1.9	215	4	US-08-252-991A-25666	Sequence 25666, A
34	7	1.9	216	3	US-08-506-553C-23	Sequence 23, Appl
35	7	1.9	226	4	US-09-107-532A-5362	Sequence 5362, Ap
36	7	1.9	235	2	US-08-529-190B-1	Sequence 1, Appl
37	7	1.9	238	4	US-09-219-983A-12	Sequence 12, Appl
38	7	1.9	240	1	US-08-414-926A-10	Sequence 10, Appl
39	7	1.9	240	2	US-08-926-922-10	Sequence 10, Appl
40	7	1.9	240	3	US-09-253-682-10	Sequence 10, Appl
41	7	1.9	240	3	US-09-527-657-10	Sequence 10, Appl
42	7	1.9	242	4	US-09-252-991A-22027	Sequence 22027, A
43	7	1.9	254	4	US-09-266-965-123	Sequence 123, Appl
44	7	1.9	267	4	US-09-219-983A-11	Sequence 11, Appl
45	7	1.9	269	4	US-08-252-991A-16773	Sequence 16773, A
46	7	1.9	276	3	US-08-506-553C-26	Sequence 26, Appl
47	7	1.9	281	4	US-09-252-991A-19858	Sequence 19858, A
48	7	1.9	307	4	US-09-252-991A-26854	Sequence 26854, A
49	7	1.9	313	4	US-09-252-991A-22411	Sequence 22411, A
50	7	1.9	322	4	US-09-252-991A-22018	Sequence 22018, A
51	7	1.9	323	4	US-09-252-991A-33113	Sequence 33113, A
52	7	1.9	340	4	US-09-328-352-5177	Sequence 5177, Ap
53	7	1.9	352	4	US-09-252-991A-17906	Sequence 17906, A
54	7	1.9	366	4	US-09-328-352-5630	Sequence 23504, A
55	7	1.9	372	4	US-09-328-352-5630	Sequence 17906, A
56	7	1.9	441	4	US-09-252-991A-24396	Sequence 24396, A
57	7	1.9	486	2	US-08-821-355A-8	Sequence 8, Appl
58	7	1.9	486	2	US-09-003-687A-8	Sequence 8, Appl
59	7	1.9	486	3	US-09-136-605-8	Sequence 8, Appl
60	7	1.9	511	2	US-08-821-355A-9	Sequence 9, Appl
61	7	1.9	511	2	US-09-003-687A-9	Sequence 9, Appl
62	7	1.9	511	3	US-08-136-605-9	Sequence 9, Appl
63	7	1.9	521	4	US-09-252-991A-18266	Sequence 18266, A
64	7	1.9	539	3	US-08-687-590-27	Sequence 27, Appl
65	7	1.9	539	4	US-09-702-705-326	Sequence 326, App
66	7	1.9	539	4	US-09-736-457-326	Sequence 326, App
67	7	1.9	546	2	US-09-067-351-1	Sequence 1, Appl
68	7	1.9	546	3	US-09-360-490-1	Sequence 1, Appl
69	7	1.9	590	3	US-09-413-814-89	Sequence 89, Appl
70	7	1.9	591	3	US-09-413-814-76	Sequence 76, Appl
71	7	1.9	609	4	US-09-252-991A-20134	Sequence 20134, A
72	7	1.9	612	4	US-08-894-454-163	Sequence 163, App
73	7	1.9	615	4	US-09-328-352-4474	Sequence 4474, Ap
74	7	1.9	627	4	US-09-345-473E-46	Sequence 46, Appl
75	7	1.9	627	4	US-09-345-473E-47	Sequence 47, Appl
76	7	1.9	628	4	US-09-345-473E-48	Sequence 48, Appl
77	7	1.9	641	4	US-09-249-585A-3	Sequence 3, Appl
78	7	1.9	641	4	US-09-410-399-4	Sequence 4, Appl
79	7	1.9	675	1	US-08-386-495-10	Sequence 10, Appl
80	7	1.9	675	5	PCT-US96-02331-10	Sequence 10, Appl
81	7	1.9	698	4	US-09-107-532A-5685	Sequence 5685, Ap
82	7	1.9	728	4	US-09-252-991A-31891	Sequence 31891, A
83	7	1.9	753	4	US-09-252-991A-17612	Sequence 17612, A
84	7	1.9	800	1	US-08-785-052-4	Sequence 4, Appl
85	7	1.9	800	2	US-08-913-581-4	Sequence 4, Appl
86	7	1.9	801	1	US-07-906-349A-6	Sequence 6, Appl
87	7	1.9	887	4	US-09-077-940A-2	Sequence 2, Appl
88	7	1.9	924	1	US-08-481-130-28	Sequence 28, Appl
89	7	1.9	924	1	US-08-656-984A-28	Sequence 28, Appl
90	7	1.9	924	1	US-08-485-604-28	Sequence 28, Appl
91	7	1.9	924	2	US-08-487-595-28	Sequence 28, Appl
92	7	1.9	971	2	US-08-724-354D-22	Sequence 22, Appl
93	7	1.9	971	3	US-09-270-984A-22	Sequence 22, Appl
94	7	1.9	971	3	US-09-177-431-8	Sequence 8, Appl
95	7	1.9	988	4	US-09-252-991A-29699	Sequence 29699, A
96	7	1.9	1000	4	US-09-252-991A-31361	Sequence 31361, A
97	7	1.9	1075	4	US-09-252-991A-18387	Sequence 18387, A
98	7	1.9	1093	5	PCT-US94-04496-55	Sequence 55, Appl
99	7	1.9	1093	5	PCT-US94-04496-55	Sequence 55, Appl
100	7	1.9	1323	1	US-08-026-138E-4	Sequence 4, Appl

ALIGNMENTS

```
RESULT 1
US-09-413-574-4
; Sequence 4, Application US/09413574
; Patent No. 6235972
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; APPLICANT: Tagliani, Laura
; TITLE OF INVENTION: Maize Rad23 Genes and Uses Thereof
; FILE REFERENCE: 0964
; CURRENT APPLICATION NUMBER: US/09/413,574
; CURRENT FILING DATE: 1999-10-06
; EARLIER APPLICATION NUMBER: 60/109,728
; EARLIER FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Zea mays
US-09-413-574-4

Query Match
Best Local Similarity 100.0%; Score 368; DB 3; Length 368;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLTVKTLKGTHTFEIRVQPNNDTIMAVKKNIEETQKDSYPMGQQLIFNGKVLKDESTLE 60
DB 1 MKLTVKTLKGTHTFEIRVQPNNDTIMAVKKNIEETQKDSYPMGQQLIFNGKVLKDESTLE 60

QY 61 ENKVNEDGFLVVMLSKGKTSGTSTSSQHSNTPATRQAPPLPAPQAPQPPVAPITTSQ 120
DB 61 ENKVNEDGFLVVMLSKGKTSGTSTSSQHSNTPATRQAPPLPAPQAPQPPVAPITTSQ 120

QY 121 PEGLPAAQPNTHNAAASNLSSGRNVDITINQLMEMGGGWDKDKVORALRAAYNNPERAV 180
DB 121 PEGLPAAQPNTHNAAASNLSSGRNVDITINQLMEMGGGWDKDKVORALRAAYNNPERAV 180

QY 181 EYLYSGIPVTAETAVPIGGGANTTDRAPTGEAGLSGIPNTAPLDLFPQASNAGGAGG 240
DB 181 EYLYSGIPVTAETAVPIGGGANTTDRAPTGEAGLSGIPNTAPLDLFPQASNAGGAGG 240

QY 241 GPLDPLRNNDPQFQAVREMVHTNPQILQPMVLVELSKONPQILRLIEENHDEFLLQNEPPE 300
DB 241 GPLDPLRNNDPQFQAVREMVHTNPQILQPMVLVELSKONPQILRLIEENHDEFLLQNEPPE 300

QY 301 GREGDFLDQPEDEMPHAI SVTPEQEAIGRLSEMGFDRARVIEAFACDRNEELAAANYL 360
DB 301 GREGDFLDQPEDEMPHAI SVTPEQEAIGRLSEMGFDRARVIEAFACDRNEELAAANYL 360

QY 361 LEHAGEED 368
DB 361 LEHAGEED 368

RESULT 2
US-09-413-574-2
; Sequence 2, Application US/09413574
; Patent No. 6235972
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; APPLICANT: Tagliani, Laura
; TITLE OF INVENTION: Maize Rad23 Genes and Uses Thereof
; FILE REFERENCE: 0964
; CURRENT APPLICATION NUMBER: US/09/413,574
; CURRENT FILING DATE: 1999-10-06
; EARLIER APPLICATION NUMBER: 60/109,728
; EARLIER FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 5
```

```
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Zea mays
US-09-413-574-2

Query Match
Best Local Similarity 3.5%; Score 13; DB 3; Length 405;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 RALRAAYNNPERA 179
DB 162 RALRAAYNNPERA 194

RESULT 3
US-09-540-014-9
; Sequence 9, Application US/09540014
; Patent No. 6380372
; GENERAL INFORMATION:
; APPLICANT: Cho, Myeong-Je
; APPLICANT: Del Val, Greg
; APPLICANT: Caillaud, Maxime
; APPLICANT: Lemauz, Peggy G.
; APPLICANT: Buchanan, Bob B.
; TITLE OF INVENTION: Barley Gene for Thioredoxin and
; TITLE OF INVENTION: NADP-Thioredoxin Reductase
; FILE REFERENCE: 2001-0701.30
; CURRENT APPLICATION NUMBER: US/09/540,014
; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 60/127,198
; PRIOR FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: US 60/169,162
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 60/177,740
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,739
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Hordeum vulgare
US-09-540-014-9

Query Match
Best Local Similarity 2.2%; Score 8; DB 4; Length 332;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 GGAGGGPL 243
DB 221 GGAGGGPL 228

RESULT 4
US-09-252-991A-25766
; Sequence 25766, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25766
```

; LENGTH: 371
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25766

Query Match 2.2%; Score 8; DB 4; Length 371;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 PEGLPAAQ 128
Db 151 PEGLPAAQ 158

RESULT 5

US-09-252-991A-28428
; Sequence 28428, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28428
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28428

Query Match 2.2%; Score 8; DB 4; Length 379;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 NAGGGAGG 240
Db 217 NAGGGAGG 224

RESULT 6

US-08-864-038A-3
; Sequence 3, Application US/08864038A
; Patent No. 6001592
; GENERAL INFORMATION:
; APPLICANT: Kunio NAKASHIMA et al.
; TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE CDNA, VECTOR
; TITLE OF INVENTION: CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID
; TITLE OF INVENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING
; TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBODY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 812-5 Hirano
; STREET: Isehinden
; CITY: Teu-city
; STATE: Mie-prefecture
; COUNTRY: JAPAN
; ZIP: 514-01
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Microsoft Windows 95
; SOFTWARE: Word Perfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/864,038A
; FILING DATE: May 28, 1997
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 8-184459
; FILING DATE: 15-July-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: C. Bruce Hamburg
; REGISTRATION NUMBER: 22,389
; REFERENCE/DOCKET NUMBER: F-5610
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)986-2340
; TELEFAX: (212)953-7733
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 738
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Pinctada fucata
; CELL TYPE: mantle epithelial cell
; FEATURE:
; NAME/KEY: peptide
; LOCATION: from 1 to 738
; IDENTIFICATION METHOD: E (by experiment)
US-08-864-038A-3

Query Match 2.2%; Score 8; DB 3; Length 738;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGG 241
Db 119 AGGGAGG 126

RESULT 7

US-08-900-230-3
; Sequence 3, Application US/08900230
; Patent No. 6329197
; GENERAL INFORMATION:
; APPLICANT: Bard, Jonathan A.
; TITLE OF INVENTION: DNA ENCODING GALANN GALR3 RECEPTORS AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of The Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,230
; FILING DATE: 23-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 52241-C/JPW/ADM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1417 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL: NO

; ANTI-SENSE: NO
; US-08-900-230-3

Query Match 2.2%; Score 8; DB 4; Length 1417;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGGG 241
Db 1237 AGGAGGG 1244

RESULT 8

US-08-961-308B-5
; Sequence 5, Application US/08961308B
; Patent No. 6248539
; GENERAL INFORMATION:
; APPLICANT: Ghadiri, Reza M.
; APPLICANT: Sailor, Michael J.
; APPLICANT: Motesharei, Kianoush
; APPLICANT: Lin, Shang-Yi
; APPLICANT: Dancil, Keiki-Pua S.
; TITLE OF INVENTION: A Porous Silicon-Based Optical Interferometric Sensor
; FILE REFERENCE: 4583-71578
; CURRENT APPLICATION NUMBER: US/08/961,308B
; CURRENT FILING DATE: 1997-10-30
; PRIOR APPLICATION NUMBER: 08/924,601
; PRIOR FILING DATE: 1997-09-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: polypeptide related to Epstein-Barr Virus Nuclear
; OTHER INFORMATION: Antigen
US-08-961-308B-5

Query Match 1.9%; Score 7; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGGG 240
Db 3 AGGAGGG 9

RESULT 9

US-08-506-553C-1
; Sequence 1, Application US/08506553C
; Patent No. 6120989
; GENERAL INFORMATION:
; APPLICANT: Vornhagen, Rolf; Hinderer, Walter; Sonneborn, Han-H.;
; APPLICANT: Plachter, Bodo; and Jahn, Gerhard
; TITLE OF INVENTION: ISOLATED HUMAN CYTOMEGALOVIRUS
; TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/506,553C

; FILING DATE: 07/25/95
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LEADER 202-PPF/MAS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; FRAGMENT TYPE: internal
US-08-506-553C-1

Query Match 1.9%; Score 7; DB 3; Length 57;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
Db 4 GGGAGGG 10

RESULT 10

US-09-149-476-615
; Sequence 615, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23

1	EARLIER	FILING DATE:	1997-08-22
2	EARLIER	APPLICATION NUMBER:	60/056,903
3	EARLIER	FILING DATE:	1997-08-22
4	EARLIER	APPLICATION NUMBER:	60/056,888
5	EARLIER	FILING DATE:	1997-08-22
6	EARLIER	APPLICATION NUMBER:	60/056,879
7	EARLIER	FILING DATE:	1997-08-22
8	EARLIER	APPLICATION NUMBER:	60/056,880
9	EARLIER	FILING DATE:	1997-08-22
10	EARLIER	APPLICATION NUMBER:	60/056,894
11	EARLIER	FILING DATE:	1997-08-22
12	EARLIER	APPLICATION NUMBER:	60/056,911
13	EARLIER	FILING DATE:	1997-08-22
14	EARLIER	APPLICATION NUMBER:	60/056,636
15	EARLIER	FILING DATE:	1997-08-22
16	EARLIER	APPLICATION NUMBER:	60/056,874
17	EARLIER	FILING DATE:	1997-08-22
18	EARLIER	APPLICATION NUMBER:	60/056,910
19	EARLIER	FILING DATE:	1997-08-22
20	EARLIER	APPLICATION NUMBER:	60/056,864
21	EARLIER	FILING DATE:	1997-08-22
22	EARLIER	APPLICATION NUMBER:	60/056,631
23	EARLIER	FILING DATE:	1997-08-22
24	EARLIER	APPLICATION NUMBER:	60/056,845
25	EARLIER	FILING DATE:	1997-08-22
26	EARLIER	APPLICATION NUMBER:	60/056,892
27	EARLIER	FILING DATE:	1997-08-22
28	EARLIER	APPLICATION NUMBER:	60/057,761
29	EARLIER	FILING DATE:	1997-08-22
30	EARLIER	APPLICATION NUMBER:	60/047,595
31	EARLIER	FILING DATE:	1997-05-23
32	EARLIER	APPLICATION NUMBER:	60/047,599
33	EARLIER	FILING DATE:	1997-05-23
34	EARLIER	APPLICATION NUMBER:	60/047,588
35	EARLIER	FILING DATE:	1997-05-23
36	EARLIER	APPLICATION NUMBER:	60/047,585
37	EARLIER	FILING DATE:	1997-05-23
38	EARLIER	APPLICATION NUMBER:	60/047,586
39	EARLIER	FILING DATE:	1997-05-23
40	EARLIER	APPLICATION NUMBER:	60/047,590
41	EARLIER	FILING DATE:	1997-05-23
42	EARLIER	APPLICATION NUMBER:	60/047,594
43	EARLIER	FILING DATE:	1997-05-23
44	EARLIER	APPLICATION NUMBER:	60/047,589
45	EARLIER	FILING DATE:	1997-05-23
46	EARLIER	APPLICATION NUMBER:	60/047,593
47	EARLIER	FILING DATE:	1997-05-23
48	EARLIER	APPLICATION NUMBER:	60/047,614
49	EARLIER	FILING DATE:	1997-05-23
50	EARLIER	APPLICATION NUMBER:	60/043,578
51	EARLIER	FILING DATE:	1997-04-11
52	EARLIER	APPLICATION NUMBER:	60/043,576
53	EARLIER	FILING DATE:	1997-04-11
54	EARLIER	APPLICATION NUMBER:	60/047,501
55	EARLIER	FILING DATE:	1997-05-23
56	EARLIER	APPLICATION NUMBER:	60/043,670
57	EARLIER	FILING DATE:	1997-04-11
58	EARLIER	APPLICATION NUMBER:	60/056,632
59	EARLIER	FILING DATE:	1997-08-22
60	EARLIER	APPLICATION NUMBER:	60/056,664
61	EARLIER	FILING DATE:	1997-08-22
62	EARLIER	APPLICATION NUMBER:	60/056,876
63	EARLIER	FILING DATE:	1997-08-22
64	EARLIER	APPLICATION NUMBER:	60/056,881
65	EARLIER	FILING DATE:	1997-08-22
66	EARLIER	APPLICATION NUMBER:	60/056,909
67	EARLIER	FILING DATE:	1997-08-22
68	EARLIER	APPLICATION NUMBER:	60/056,875
69	EARLIER	FILING DATE:	1997-08-22
70	EARLIER	APPLICATION NUMBER:	60/056,862
71	EARLIER	FILING DATE:	1997-08-22
72	EARLIER	APPLICATION NUMBER:	60/056,887
73	EARLIER	FILING DATE:	1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match 1.9%; Score 7; DB 4; Length 61;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 TGTSSQ 89
|||||||
Db 40 TGTSSQ 46

RESULT 11
US-09-249-542-9
; Sequence 9, Application US/09249542A
; Patent No. 6392029

; GENERAL INFORMATION:
; APPLICANT: Ludwig, Linda B.
; TITLE OF INVENTION: HIV Chemokines
; FILE REFERENCE: 11520.0122
; CURRENT APPLICATION NUMBER: US/09/249,542A
; CURRENT FILING DATE: 1999-02-12
; EARLIER APPLICATION NUMBER: US 60/074,640
; EARLIER FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 26

; SEQ ID NO 9
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus

; FEATURE:
; OTHER INFORMATION: HIV chemokine protein (2nd start site; (+1) ribosomal
; OTHER INFORMATION: shift at nucleotide 201)from SF2 HIV
US-09-249-542-9

Query Match 1.9%; Score 7; DB 4; Length 81;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 SGTSGTS 86
|||||||
Db 58 SGTSGTS 64

RESULT 12
US-09-249-542-8
; Sequence 8, Application US/09249542A
; Patent No. 6392029

; GENERAL INFORMATION:
; APPLICANT: Ludwig, Linda B.
; TITLE OF INVENTION: HIV Chemokines
; FILE REFERENCE: 11520.0122
; CURRENT APPLICATION NUMBER: US/09/249,542A
; CURRENT FILING DATE: 1999-02-12
; EARLIER APPLICATION NUMBER: US 60/074,640
; EARLIER FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 26

; SEQ ID NO 8
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus

; FEATURE:

; OTHER INFORMATION: HIV chemokine protein (2nd start site; (-1) ribosomal
; OTHER INFORMATION: shifts at nucleotides 200 and 265)from SF2 HIV
US-09-249-542-8

Query Match 1.9%; Score 7; DB 4; Length 82;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 SGTSGTS 86
|||||||
Db 59 SGTSGTS 65

RESULT 13

US-09-036-113-2
; Sequence 2, Application US/09036113
; Patent No. 6077825

; GENERAL INFORMATION:
; APPLICANT: Cupp, Mary S.
; APPLICANT: Cupp, Eddie W.
; TITLE OF INVENTION: ANTITHROMBIN PROTEIN AND DNA SEQUENCES
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: W. Murray Spruill (Alston & Bird, LLP)
; STREET: 3605 Glenwood Ave. Suite 310
; CITY: Raleigh
; STATE: NC
; COUNTRY: USA
; ZIP: 27622

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036.113
; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5721-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919 420 2202
; TELEFAX: 919 881 3175

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 96 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-09-036-113-2

Query Match 1.9%; Score 7; DB 3; Length 96;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 349 CDRNEEL 355
|||||||
Db 64 CDRNEEL 70

RESULT 14

US-09-556-166-2
; Sequence 2, Application US/09556166
; Patent No. 6465214

; GENERAL INFORMATION:
; APPLICANT: Cupp, Mary S.
; APPLICANT: Cupp, Eddie W.
; TITLE OF INVENTION: ANTITHROMBIN PROTEIN AND DNA SEQUENCES
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: W. Murray Spruill (Alston & Bird, LLP)

; STREET: 3605 Glenwood Ave. Suite 310
 ; CITY: Raleigh
 ; STATE: NC
 ; COUNTRY: USA
 ; ZIP: 27622
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/556,166
 ; FILING DATE: 21-APR-2000
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 09/036,113
 ; FILING DATE: 06-MAR-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Spruill, W. Murray
 ; REGISTRATION NUMBER: 32,943
 ; REFERENCE/DOCKET NUMBER: 5721-5
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 919 420 2202
 ; TELEFAX: 919 881 3175
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 96 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-556-166-2

Query Match 1.9%; Score 7; DB 4; Length 96;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 349 CDRNEEL 355
 Db 64 CDRNEEL 70

RESULT 15

; Sequence 29877, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 29877
 ; LENGTH: 96
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-29877

Query Match 1.9%; Score 7; DB 4; Length 96;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 STGTSSS 88
 Db 44 STGTSSS 50

RESULT 16

; US-09-249-542-16
 ; Sequence 16, Application US/09249542A
 ; Patent No. 6392029
 ; GENERAL INFORMATION:
 ; APPLICANT: Ludwig, Linda B.
 ; TITLE OF INVENTION: HIV Chemokines
 ; FILE REFERENCE: 11520.0122
 ; CURRENT APPLICATION NUMBER: US/09/249,542A
 ; CURRENT FILING DATE: 1999-02-12
 ; EARLIER APPLICATION NUMBER: US 60/074,640
 ; EARLIER FILING DATE: 1998-02-13
 ; NUMBER OF SEQ ID NOS: 26
 ; SEQ ID NO 16
 ; LENGTH: 111
 ; TYPE: PRT
 ; ORGANISM: Human immunodeficiency virus
 ; FEATURE:
 ; OTHER INFORMATION: HIV chemokine protein from YU2 strain with a (+1) and (-)
 ; OTHER INFORMATION: 1) ribosomal frameshift
 ; US-09-249-542-16

Query Match 1.9%; Score 7; DB 4; Length 111;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 SGSTGTS 86
 Db 89 SGSTGTS 95

RESULT 17

; US-09-249-542-17
 ; Sequence 17, Application US/09249542A
 ; Patent No. 6392029
 ; GENERAL INFORMATION:
 ; APPLICANT: Ludwig, Linda B.
 ; TITLE OF INVENTION: HIV Chemokines
 ; FILE REFERENCE: 11520.0122
 ; CURRENT APPLICATION NUMBER: US/09/249,542A
 ; CURRENT FILING DATE: 1999-02-12
 ; EARLIER APPLICATION NUMBER: US 60/074,640
 ; EARLIER FILING DATE: 1998-02-13
 ; NUMBER OF SEQ ID NOS: 26
 ; SEQ ID NO 17
 ; LENGTH: 111
 ; TYPE: PRT
 ; ORGANISM: Human immunodeficiency virus
 ; FEATURE:
 ; OTHER INFORMATION: HIV chemokine protein (1st start site; (-1) ribosomal
 ; OTHER INFORMATION: frameshift at nucleotide 133, (+1) ribosomal frameshift at
 ; OTHER INFORMATION: nucleotide 265) from YU2 HIV
 ; US-09-249-542-17

Query Match 1.9%; Score 7; DB 4; Length 111;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 SGSTGTS 86
 Db 89 SGSTGTS 95

RESULT 18

; US-09-249-542-18
 ; Sequence 18, Application US/09249542A
 ; Patent No. 6392029
 ; GENERAL INFORMATION:
 ; APPLICANT: Ludwig, Linda B.
 ; TITLE OF INVENTION: HIV Chemokines
 ; FILE REFERENCE: 11520.0122
 ; CURRENT APPLICATION NUMBER: US/09/249,542A
 ; CURRENT FILING DATE: 1999-02-12
 ; EARLIER APPLICATION NUMBER: US 60/074,640

```
; EARLIER FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 26
; SEQ ID NO 18
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
; FEATURE:
; NAME/KEY: Xaa is any amino acid
; LOCATION: 28,37,47
; OTHER INFORMATION: HIV chemokine protein with read-through from YU2 HIV
US-09-249-542-18

Query Match      1.9%; Score 7; DB 4; Length 111;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      80 SGGTGTS 86
Db      89 SGGTGTS 95
|||||

RESULT 19
US-09-249-542-7
; Sequence 7, Application US/09249542A
; Patent No. 6392029
; GENERAL INFORMATION:
; APPLICANT: Ludwig, Linda B.
; TITLE OF INVENTION: HIV Chemokines
; FILE REFERENCE: 11520.0122
; CURRENT APPLICATION NUMBER: US/09/249,542A
; CURRENT FILING DATE: 1999-02-12
; EARLIER APPLICATION NUMBER: US 60/074,640
; EARLIER FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 26
; SEQ ID NO 7
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
; FEATURE:
; OTHER INFORMATION: HIV chemokine protein (1st start site; (+1) ribosomal
; OTHER INFORMATION: frameshift at nucleotide 133; and (-1) ribosomal shift at
; OTHER INFORMATION: nucleotide 265) from SF2 HIV
US-09-249-542-7

Query Match      1.9%; Score 7; DB 4; Length 112;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      80 SGGTGTS 86
Db      89 SGGTGTS 95
|||||

RESULT 20
US-09-249-542-10
; Sequence 10, Application US/09249542A
; Patent No. 6392029
; GENERAL INFORMATION:
; APPLICANT: Ludwig, Linda B.
; TITLE OF INVENTION: HIV Chemokines
; FILE REFERENCE: 11520.0122
; CURRENT APPLICATION NUMBER: US/09/249,542A
; CURRENT FILING DATE: 1999-02-12
; EARLIER APPLICATION NUMBER: US 60/074,640
; EARLIER FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 26
; SEQ ID NO 10
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
; FEATURE:
; OTHER INFORMATION: HIV chemokine protein (1st start site at nucleotide 52;
; OTHER INFORMATION: (+1) ribosomal frameshift at nucleotide 133, (-1) ribosomal
; OTHER INFORMATION: frameshift at nucleotide 265) from SF2 HIV
US-09-249-542-10

Query Match      1.9%; Score 7; DB 4; Length 112;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      80 SGGTGTS 86
Db      89 SGGTGTS 95
|||||

RESULT 21
US-09-249-542-11
; Sequence 11, Application US/09249542A
; Patent No. 6392029
; GENERAL INFORMATION:
; APPLICANT: Ludwig, Linda B.
; TITLE OF INVENTION: HIV Chemokines
; FILE REFERENCE: 11520.0122
; CURRENT APPLICATION NUMBER: US/09/249,542A
; CURRENT FILING DATE: 1999-02-12
; EARLIER APPLICATION NUMBER: US 60/074,640
; EARLIER FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 26
; SEQ ID NO 11
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
; FEATURE:
; OTHER INFORMATION: HIV chemokine protein (1st start site; (-1) ribosomal
; OTHER INFORMATION: frameshift at nucleotide 133, (+1) ribosomal frameshift at
; OTHER INFORMATION: nucleotide 265), from (Pt)CNS HIV
US-09-249-542-11

Query Match      1.9%; Score 7; DB 4; Length 112;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      80 SGGTGTS 86
Db      89 SGGTGTS 95
|||||

RESULT 22
US-09-249-542-14
; Sequence 14, Application US/09249542A
; Patent No. 6392029
; GENERAL INFORMATION:
; APPLICANT: Ludwig, Linda B.
; TITLE OF INVENTION: HIV Chemokines
; FILE REFERENCE: 11520.0122
; CURRENT APPLICATION NUMBER: US/09/249,542A
; CURRENT FILING DATE: 1999-02-12
; EARLIER APPLICATION NUMBER: US 60/074,640
; EARLIER FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 26
; SEQ ID NO 14
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
; FEATURE:
; OTHER INFORMATION: HIV chemokine protein (1st start site; (-1) ribosomal
; OTHER INFORMATION: frameshift at nucleotide 133, (+1) ribosomal frameshift at
; OTHER INFORMATION: nucleotide 265) from SF2 HIV
US-09-249-542-14

Query Match      1.9%; Score 7; DB 4; Length 112;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      80 SGGTGTS 86
Db      89 SGGTGTS 95
|||||
```

Db 89 SGSTGTS 95

RESULT 23

US-09-249-542-15
; Sequence 15, Application US/09249542A
; Patent No. 6392029
; GENERAL INFORMATION:
; APPLICANT: Ludwig, Linda B.
; TITLE OF INVENTION: HIV Chemokines
; FILE REFERENCE: 11520.0122
; CURRENT APPLICATION NUMBER: US/09/249,542A
; CURRENT FILING DATE: 1999-02-12
; EARLIER APPLICATION NUMBER: US 60/074,640
; EARLIER FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 26
; SEQ ID NO 15
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
; FEATURE:
; NAME/KEY: Xaa is any amino acid
; LOCATION: 28,37,47
; OTHER INFORMATION: HIV chemokine protein with read-through from SF2 HIV
US-09-249-542-15

Query Match 1.9%; Score 7; DB 4; Length 112;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 SGSTGTS 86

Db 89 SGSTGTS 95

RESULT 24

US-09-249-542-19
; Sequence 19, Application US/09249542A
; Patent No. 6392029
; GENERAL INFORMATION:
; APPLICANT: Ludwig, Linda B.
; TITLE OF INVENTION: HIV Chemokines
; FILE REFERENCE: 11520.0122
; CURRENT APPLICATION NUMBER: US/09/249,542A
; CURRENT FILING DATE: 1999-02-12
; EARLIER APPLICATION NUMBER: US 60/074,640
; EARLIER FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 26
; SEQ ID NO 19
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
; FEATURE:
; NAME/KEY: 28,37,97
; LOCATION: Xaa is any amino acid
; OTHER INFORMATION: HIV chemokine protein with read-through from ELI strain
; OTHER INFORMATION: of HIV
US-09-249-542-19

Query Match 1.9%; Score 7; DB 4; Length 112;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 GSTGTSS 87

Db 90 GSTGTSS 96

RESULT 25

US-09-249-542-24
; Sequence 24, Application US/09249542A
; Patent No. 6392029
; GENERAL INFORMATION:

; APPLICANT: Ludwig, Linda B.
; TITLE OF INVENTION: HIV Chemokines
; FILE REFERENCE: 11520.0122
; CURRENT APPLICATION NUMBER: US/09/249,542A
; CURRENT FILING DATE: 1999-02-12
; EARLIER APPLICATION NUMBER: US 60/074,640
; EARLIER FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 26
; SEQ ID NO 24
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
; FEATURE:
; NAME/KEY: Xaa is any amino acid
; LOCATION: 97
; OTHER INFORMATION: HIV chemokine protein with (-1) frameshift at nucleotide
; OTHER INFORMATION: 133, (+1) frameshift at nucleotide 205, readthrough at
; OTHER INFORMATION: nucleotides 341-343 from pHIVCAT
US-09-249-542-24

Query Match 1.9%; Score 7; DB 4; Length 112;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 SGSTGTS 86

Db 89 SGSTGTS 95

RESULT 26

US-09-996-243-420
; Sequence 420, Application US/09996243
; Patent No. 6478825
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C13
; CURRENT APPLICATION NUMBER: US/09/996,243
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24

;	PRIOR FILING DATE: 1998-06-17
;	PRIOR APPLICATION NUMBER: 60/089598
;	PRIOR FILING DATE: 1998-06-17
;	PRIOR APPLICATION NUMBER: 60/089599
;	PRIOR FILING DATE: 1998-06-17
;	PRIOR APPLICATION NUMBER: 60/089600
;	PRIOR FILING DATE: 1998-06-17
;	PRIOR APPLICATION NUMBER: 60/089653
;	PRIOR FILING DATE: 1998-06-17
;	PRIOR APPLICATION NUMBER: 60/089801
;	PRIOR FILING DATE: 1998-06-18
;	PRIOR APPLICATION NUMBER: 60/089907
;	PRIOR FILING DATE: 1998-06-18
;	PRIOR APPLICATION NUMBER: 60/089908
;	PRIOR FILING DATE: 1998-06-18
;	PRIOR APPLICATION NUMBER: 60/089947
;	PRIOR FILING DATE: 1998-06-19
;	PRIOR APPLICATION NUMBER: 60/089948
;	PRIOR FILING DATE: 1998-06-19
;	PRIOR APPLICATION NUMBER: 60/089952
;	PRIOR FILING DATE: 1998-06-19
;	PRIOR APPLICATION NUMBER: 60/090246
;	PRIOR FILING DATE: 1998-06-22
;	PRIOR APPLICATION NUMBER: 60/090252
;	PRIOR FILING DATE: 1998-06-22
;	PRIOR APPLICATION NUMBER: 60/090254
;	PRIOR FILING DATE: 1998-06-22
;	PRIOR APPLICATION NUMBER: 60/090349
;	PRIOR FILING DATE: 1998-06-23
;	PRIOR APPLICATION NUMBER: 60/090355
;	PRIOR FILING DATE: 1998-06-23
;	PRIOR APPLICATION NUMBER: 60/090429
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090431
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090435
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090444
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090445
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090472
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090535
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090540
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090542
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090557
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090676
;	PRIOR FILING DATE: 1998-06-25
;	PRIOR APPLICATION NUMBER: 60/090678
;	PRIOR FILING DATE: 1998-06-25
;	PRIOR APPLICATION NUMBER: 60/090690
;	PRIOR FILING DATE: 1998-06-25
;	PRIOR APPLICATION NUMBER: 60/090696
;	PRIOR FILING DATE: 1998-06-25
;	PRIOR APPLICATION NUMBER: 60/090862
;	PRIOR FILING DATE: 1998-06-26
;	PRIOR APPLICATION NUMBER: 60/090863
;	PRIOR FILING DATE: 1998-06-26
;	PRIOR APPLICATION NUMBER: 60/091360
;	PRIOR FILING DATE: 1998-07-01
;	PRIOR APPLICATION NUMBER: 60/091478
;	PRIOR FILING DATE: 1998-07-02
;	PRIOR APPLICATION NUMBER: 60/091544
;	PRIOR FILING DATE: 1998-07-01

; PRIOR APPLICATION NUMBER: 60/091519
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091626
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091633
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091978
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/091982
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/092182
 ; PRIOR FILING DATE: 1998-07-09

Query Match 1.9%; Score 7; DB 4; Length 128;
 Best Local Similarity 100.0%; Pred. No. 62;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGG 241
 DB 83 GGGAGG 89

RESULT 27

US-09-328-352-6723
 ; Sequence 6723, Application US/09328352
 ; Patent No. 6562958

GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 6723
 ; LENGTH: 160
 ; TYPE: PRT

; ORGANISM: Acinetobacter baumannii
 US-09-328-352-6723

Query Match 1.9%; Score 7; DB 4; Length 160;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 237 GAGGGPL 243
 DB 128 GAGGGPL 134

RESULT 28

US-08-403-852D-30
 ; Sequence 30, Application US/08403852D
 ; Patent No. 5891695

GENERAL INFORMATION:

; APPLICANT: Blanc, Veronique
 ; APPLICANT: Blanche, Francis
 ; APPLICANT: Crouzet, Joel
 ; APPLICANT: Jacques, Nathalie
 ; APPLICANT: Lacroix, Patricia
 ; APPLICANT: Thibaut, Denis
 ; APPLICANT: Zagorec, Monique
 ; APPLICANT: Debusche, Laurent
 ; APPLICANT: De Crecy-Lagard, Valerie
 ; TITLE OF INVENTION: Polypeptides Involved In The
 ; TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences
 ; TITLE OF INVENTION: Coding For These Polypeptides And Their Use
 ; NUMBER OF SEQUENCES: 43
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
 ; STREET: 1300 I Street, N.W., Suite 700
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA

; ZIP: 20005-3315
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/403.852D
 ; FILING DATE: 10-MAY-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/FR 93/00923
 ; FILING DATE: 25-SEP-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: FR 92/11441
 ; FILING DATE: 25-SEP-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Meyers, Kenneth J.
 ; REGISTRATION NUMBER: 25,146
 ; REFERENCE/DOCKET NUMBER: 03806.0054-00000
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 408-4000
 ; TELEFAX: (202) 408-4400
 ; INFORMATION FOR SEQ ID NO: 30:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 161 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 US-08-403-852D-30

Query Match 1.9%; Score 7; DB 2; Length 161;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 TSGSTGT 85
 DB 81 TSGSTGT 87

RESULT 29

US-08-510-646B-31
 ; Sequence 31, Application US/08510646B
 ; Patent No. 6077699

GENERAL INFORMATION:

; APPLICANT: Blanc, Veronique
 ; APPLICANT: Blanche, Francis
 ; APPLICANT: Crouzet, Joel
 ; APPLICANT: Jacques, Nathalie
 ; APPLICANT: Lacroix, Patricia
 ; APPLICANT: Thibaut, Denis
 ; APPLICANT: Zagorec, Monique
 ; APPLICANT: Debusche, Laurent
 ; APPLICANT: De Crecy-Lagard, Valerie
 ; TITLE OF INVENTION: Polypeptides Involved In The
 ; TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences
 ; TITLE OF INVENTION: Coding For These Polypeptides And Their Use
 ; NUMBER OF SEQUENCES: 45
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
 ; STREET: 1300 I Street, N.W., Suite 700
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA

; ZIP: 20005-3315

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/510.646B
 ; FILING DATE: 03-AUG-1995
 ; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/403,852
;; FILING DATE: 10-MAY-1995
;; PRIOR APPLICATION DATA: PCT/FR 93/00923
;; APPLICATION NUMBER: PCT/FR 93/00923
;; FILING DATE: 25-SEP-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: FR 92/11441
;; FILING DATE: 25-SEP-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Meyers, Kenneth J.
;; REGISTRATION NUMBER: 25,146
;; REFERENCE/DOCKET NUMBER: 03806.0054-01000
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 408-4400
;; TELEFAX: (202) 408-4400
;; INFORMATION FOR SEQ ID NO: 31:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 161 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; US-08-510-646B-31

Query Match 1.9%; Score 7; DB 3; Length 161;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 TSGSTGT 85
Db 81 TSGSTGT 87

RESULT 30
US-09-231-818-30
; Sequence 30, Application US/09231818
; Patent No. 6171846
; GENERAL INFORMATION:
; APPLICANT: Blanc, Veronique
; APPLICANT: Blanche, Francis
; APPLICANT: Crouzet, Joel
; APPLICANT: Jacques, Nathalie
; APPLICANT: Lacroix, Patricia
; APPLICANT: Thibaut, Denis
; APPLICANT: Zagorec, Monique
; APPLICANT: Debussche, Laurent
; APPLICANT: De Crecy-Lagard, Valerie
; TITLE OF INVENTION: Polypeptides Involved In The
; TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences
; TITLE OF INVENTION: Coding For These Polypeptides And Their Use
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/231,818
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,852
; FILING DATE: 10-MAY-1995
; APPLICATION NUMBER: PCT/FR 93/00923
; FILING DATE: 25-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/11441

;; FILING DATE: 25-SEP-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Meyers, Kenneth J.
;; REGISTRATION NUMBER: 25,146
;; REFERENCE/DOCKET NUMBER: 03806.0054-00000
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 408-4000
;; TELEFAX: (202) 408-4400
;; INFORMATION FOR SEQ ID NO: 30:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 161 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; US-09-231-818-30

Query Match 1.9%; Score 7; DB 3; Length 161;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 TSGSTGT 85
Db 81 TSGSTGT 87

RESULT 31
US-09-562-737-112
; Sequence 112, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 112
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
; US-09-562-737-112

Query Match 1.9%; Score 7; DB 4; Length 185;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KLTVKTL 8
Db 71 KLTVKTL 77

RESULT 32
US-08-606-143-40
; Sequence 40, Application US/08606143
; Patent No. 5856155
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; TITLE OF INVENTION: COMPOUNDS AND RELATED METHODS FOR
; TITLE OF INVENTION: MODULATING POTASSIUM ION CHANNELS AND ASSAYS FOR SUCH
; TITLE OF INVENTION: COMPOUNDS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL USA
; COUNTRY: USA
; ZIP: 60601

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/606,143
 ; FILING DATE: 23-FEB-1996
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kilyk Jr., John
 ; REGISTRATION NUMBER: 30763
 ; REFERENCE/DOCKET NUMBER: 71756
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (312) 616-5600
 ; TELEFAX: (312) 616-5700
 ; TELEX: 25-3533
 ; INFORMATION FOR SEQ ID NO: 40:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 208 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-606-143-40

Query Match 1.9%; Score 7; DB 2; Length 208;
 Best Local Similarity 100.0%; Pred. No. 95;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGAGG 240
 |||||
 Db 163 AGGAGG 169

RESULT 33

; US-09-252-991A-25666
 ; Sequence 25666, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 25666
 ; LENGTH: 215
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-25666

Query Match 1.9%; Score 7; DB 4; Length 215;
 Best Local Similarity 100.0%; Pred. No. 98;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 GGGAGG 241
 |||||
 Db 136 GGGAGG 142

RESULT 34

; US-08-506-553C-23
 ; Sequence 23, Application US/08506553C
 ; Patent No. 6120989
 ; GENERAL INFORMATION:
 ; APPLICANT: Vornhagen, Rolf; Hinderer, Walter; Sonneborn, Han-H.;
 ; APPLICANT: Plachter, Bodo; and Jahn, Gerhard

; TITLE OF INVENTION: ISOLATED HUMAN CYTOMEGALOVIRUS
 ; TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF
 ; NUMBER OF SEQUENCES: 26
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Felfe & Lynch
 ; STREET: 805 Third Avenue
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10022
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: PC-DOS
 ; SOFTWARE: Wordperfect
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/506,553C
 ; FILING DATE: 07/25/95
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mary Anne Schofield
 ; REGISTRATION NUMBER: 36,669
 ; REFERENCE/DOCKET NUMBER: LEDE 202-PFF/MAS
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 688-9200
 ; TELEFAX: (212) 838-3884
 ; INFORMATION FOR SEQ ID NO: 23:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 216 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: Polypeptide
 ; FRAGMENT TYPE: internal
 ; US-08-506-553C-23

Query Match 1.9%; Score 7; DB 3; Length 216;
 Best Local Similarity 100.0%; Pred. No. 99;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 GGGAGG 241
 |||||
 Db 163 GGGAGG 169

RESULT 35

; US-09-107-532A-5362
 ; Sequence 5362, Application US/09107532A
 ; Patent No. 6583275
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 7310
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 ; STREET: 100 Beaver Street
 ; CITY: Waltham
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02154
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: CD-ROM ISO9660
 ; COMPUTER: PC
 ; OPERATING SYSTEM: <Unknown>
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/107,532A
 ; FILING DATE: 30-Jun-1998
 ; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-5277
INFORMATION FOR SEQ ID NO: 5362:
SEQUENCE CHARACTERISTICS:
LENGTH: 226 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...226
SEQUENCE DESCRIPTION: SEQ ID NO: 5362:
US-09-107-532A-5362

Query Match 1.9%; Score 7; DB 4; Length 226;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 SNLSSGR 143
|||||
DB 53 SNLSSGR 59

RESULT 36

US-08-529-190B-1
; Sequence 1, Application US/08529190B
; Patent No. 5833991
; GENERAL INFORMATION:
; APPLICANT: Masucci, Maria G.
; TITLE OF INVENTION: GLYCINE-CONTAINING SEQUENCES
; TITLE OF INVENTION: CONFERRING INVISIBILITY TO THE IMMUNE SYSTEM
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: One Financial Center
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Wordperfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,190B
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE9501324-9
; FILING DATE: 10-APR-1995
; PRIOR APPLICATION DATA: US08/522,595
; APPLICATION NUMBER: 40,489
; FILING DATE: 01-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Ph.D., Kathleen A
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: 3255/53015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-345-9100
; TELEFAX: 617-345-9111

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 235 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-529-190B-1

Query Match 1.9%; Score 7; DB 2; Length 235;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGG 240
|||||
DB 22 AGGAGG 28

RESULT 37

US-09-219-983A-12
; Sequence 12, Application US/09219983A
; Patent No. 6380159
; GENERAL INFORMATION:
; APPLICANT: Wolfner, Mariana
; APPLICANT: Lung, Oliver
; APPLICANT: Tram, Khanh-Uyen
; TITLE OF INVENTION: GENES FOR MALE ACCESSORY GLAND PROTEINS IN DROSOPHILA
; TITLE OF INVENTION: MELANOGASTER
; FILE REFERENCE: 19603/1791
; CURRENT APPLICATION NUMBER: US/09/219,983A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/071,315
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-219-983A-12

Query Match 1.9%; Score 7; DB 4; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGG 241
|||||
DB 22 GGGAGG 28

RESULT 38

US-08-414-926A-10
; Sequence 10, Application US/08414926A
; Patent No. 5721354
; GENERAL INFORMATION:
; APPLICANT: Spaete, Richard
; APPLICANT: Cha, Tai-An
; TITLE OF INVENTION: NOVEL HUMAN CYTOMEGALOVIRUS
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
; STREET: 5 Palo Alto Square
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-2155
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/414,926A

; FILING DATE: March 31, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: AVIR-011/OOUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-494-7622
; TELEFAX: 415-857-0663
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: tol.04
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..240
; OTHER INFORMATION: /label= UL136
; US-08-414-926A-10

Query Match 1.9%; Score 7; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGG 240
|||
Db 186 AGGGAGG 192

RESULT 39

US-08-926-922-10
; Sequence 10, Application US/08926922
; Patent No. 5925751
; GENERAL INFORMATION:
; APPLICANT: Spaete, Richard
; APPLICANT: Cha, Tai-An
; TITLE OF INVENTION: NOVEL HUMAN CYTOMEGALOVIRUS
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Luann Cserr Attorney at Law
; STREET: 750 Arimo Avenue
; CITY: Oakland
; STATE: CA
; COUNTRY: USA
; ZIP: 94610
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/926,922
; FILING DATE: September 10, 1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: AVIR 11A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-834-1448
; TELEFAX: 510-839-7810
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: tol.04
; FEATURE:

; NAME/KEY: Protein
; LOCATION: 1..240
; OTHER INFORMATION: /label= UL136
; US-08-926-922-10

Query Match 1.9%; Score 7; DB 2; Length 240;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGG 240
|||
Db 186 AGGGAGG 192

RESULT 40

US-09-253-682-10
; Sequence 10, Application US/09253682
; Patent No. 6040170
; GENERAL INFORMATION:
; APPLICANT: Spaete, Richard
; APPLICANT: Cha, Tai-An
; TITLE OF INVENTION: NOVEL HUMAN CYTOMEGALOVIRUS
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Luann Cserr Attorney at Law
; STREET: 750 Arimo Avenue
; CITY: Oakland
; STATE: CA
; COUNTRY: USA
; ZIP: 94610
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/253,682
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/926,922
; FILING DATE: September 10, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: AVIR 11A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-834-1448
; TELEFAX: 510-839-7810
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: tol.04
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..240
; OTHER INFORMATION: /label= UL136
; US-09-253-682-10

Query Match 1.9%; Score 7; DB 3; Length 240;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 AGGGAGG 240
|||
Db 186 AGGGAGG 192

RESULT 41

```

US-09-527-657-10
; Sequence 10, Application US/09527657
; Patent No. 6291236
; GENERAL INFORMATION:
; APPLICANT: Spaete, Richard
; Cha, Tai-An
; TITLE OF INVENTION: NOVEL HUMAN CYTOMEGALOVIRUS
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Luann Cserr Attorney at Law
; STREET: 750 Arimo Avenue
; CITY: Oakland
; STATE: CA
; COUNTRY: USA
; ZIP: 94610
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/527,657
; FILING DATE: 17-Mar-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/926,922
; FILING DATE: September 10, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: AVIR 11A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-834-1448
; TELEFAX: 510-839-7810
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: tol.04
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..240
; OTHER INFORMATION: /label= UL136
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-527-657-10

Query Match 1.9%; Score 7; DB 3; Length 240;
Best Local Similarity 100.0%; Pred.No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGGAGG 240
Db 186 AGGGAGG 192
|||||

RESULT 42
US-09-252-991A-22027
; Sequence 22027, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; BEST LOCAL SIMILARITY: 100.0%
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

```

```

; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22027
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22027

Query Match      1.9%; Score 7; DB 4; Length 242;
Best Local Similarity 100.0%; Pred.No.1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels

Oy      235 GGGAGGG 241
          |||||
Db      44 GGGAGGG 50

RESULT 43
US-09-266-965-123
; Sequence 123, Application US/09266965
; Patent NO. 6495348
; GENERAL INFORMATION:
; APPLICANT: Sherman, D
; APPLICANT: Mao, Y
; APPLICANT: Varoglu, M
; APPLICANT: He, M
; APPLICANT: Sheldon, P
; TITLE OF INVENTION: Mitomycin biosynthetic gene cluster
; FILE REFERENCE: 600.456US1
; CURRENT APPLICATION NUMBER: US/09/266,965
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: US 08/624,447
; EARLIER FILING DATE: 1996-08-19
; EARLIER APPLICATION NUMBER: PCT/US94/11279
; EARLIER FILING DATE: 1994-10-06
; EARLIER APPLICATION NUMBER: US 08/133,963
; EARLIER FILING DATE: 1993-10-07
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 123
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Streptomyces lavendulae
US-09-266-965-123

Query Match      1.9%; Score 7; DB 4; Length 254;
Best Local Similarity 100.0%; Pred.No.1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels

Oy      201 GANTTDR 207
          |||||
Db      193 GANTTDR 199

RESULT 44
US-09-219-983A-11
; Sequence 11, Application US/09219983A
; Patent No. 6380159
; GENERAL INFORMATION:
; APPLICANT: Wolfner, Mariana
; APPLICANT: Lung, Oliver
; APPLICANT: Tram, Khanh-Uyen
; TITLE OF INVENTION: GENES FOR MALE ACCESSORY GLAND PROTEIN
; FILE REFERENCE: 19603/1791
; CURRENT APPLICATION NUMBER: US/09/219,983A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/071,315
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 267
; TYPE: PRT

```

```
; ORGANISM: Drosophila melanogaster
US-09-219-983A-11

Query Match      1.9%; Score 7; DB 4; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      235 GGGAGGG 241
Db      50 GGGAGGG 56
|||||

RESULT 45
US-09-252-991A-16773
; Sequence 16773, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16773
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16773

Query Match      1.9%; Score 7; DB 4; Length 269;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      234 AGGGAGG 240
Db      111 AGGGAGG 117
|||||

RESULT 46
US-08-506-553C-26
; Sequence 26, Application US/08506553C
; Patent No. 6120989
; GENERAL INFORMATION:
; APPLICANT: Vornhagen, Rolf; Hinderer, Walter; Sonneborn, Han-H.;
; APPLICANT: Plachter, Bodo; and Jahn, Gerhard
; TITLE OF INVENTION: ISOLATED HUMAN CYTOMEGALOVIRUS
; TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/506,553C
; FILING DATE: 07/25/95
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LEADER 202-PFF/MAS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 276 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Polypeptide
; FRAGMENT TYPE: internal
US-08-506-553C-26

Query Match      1.9%; Score 7; DB 3; Length 276;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      235 GGGAGGG 241
Db      163 GGGAGGG 169
|||||

RESULT 47
US-09-252-991A-19858
; Sequence 19858, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19858
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19858

Query Match      1.9%; Score 7; DB 4; Length 281;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      120 QPGLPLA 126
Db      262 QPGLPLA 268
|||||

RESULT 48
US-09-252-991A-26854
; Sequence 26854, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26854
```

; LENGTH: 307
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26854

Query Match 1.9%; Score 7; DB 4; Length 307;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 POPVPAP 115
|||||
Db 151 POPVPAP 157

RESULT 49

US-09-252-991A-22411
; Sequence 22411, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 22411

; LENGTH: 313

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-22411

Query Match 1.9%; Score 7; DB 4; Length 313;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
|||||
Db 102 GGGAGGG 108

RESULT 50

US-09-252-991A-22018
; Sequence 22018, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 22018

; LENGTH: 322

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-22018

Query Match 1.9%; Score 7; DB 4; Length 322;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 298 PFEQEGG 304
|||||

Db 15 PFEQEGG 21

RESULT 51

US-09-252-991A-33113
; Sequence 33113, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 33113

; LENGTH: 323

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-33113

Query Match 1.9%; Score 7; DB 4; Length 323;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 HSNTPAT 96
|||||
Db 55 HSNTPAT 61

RESULT 52

US-09-328-352-5177
; Sequence 5177, Application US/09328352
; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; FILE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 5177

; LENGTH: 340

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii

US-09-328-352-5177

Query Match 1.9%; Score 7; DB 4; Length 340;

Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 EFLQLLN 296
|||||
Db 133 EFLQLLN 139

RESULT 53

US-09-252-991A-17906
; Sequence 17906, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17906
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17906

Query Match 1.9%; Score 7; DB 4; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AGGAGG 240
Db 313 AGGAGG 319

RESULT 54
US-09-252-991A-23504
; Sequence 23504, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23504
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23504

Query Match 1.9%; Score 7; DB 4; Length 366;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 GGAGGGP 242
Db 192 GGAGGGP 198

RESULT 55
US-09-328-352-5630
; Sequence 5630, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5630
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5630

Query Match 1.9%; Score 7; DB 4; Length 372;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 VQNDTI 23

Db 244 VQNDTI 250
|||||

RESULT 56
US-09-252-991A-24396
; Sequence 24396, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24396
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24396

Query Match 1.9%; Score 7; DB 4; Length 441;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 GGAGGGP 242
Db 188 GGAGGGP 194

RESULT 57
US-08-821-355A-8
; Sequence 8, Application US/08821355A
; Patent No. 5851775
; GENERAL INFORMATION:

; APPLICANT: Barker, Nick
; APPLICANT: Clevers, Hans
; APPLICANT: Korinek, Vladimir
; APPLICANT: Morin, Patrice
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Vogelstein, Bert
; APPLICANT: Sparks, Andrew
; TITLE OF INVENTION: Beta Catenin, TCF-4, and APC
; TITLE OF INVENTION: Interact to Prevent Cancer
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 1001 G Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20001

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/821,355A
FILING DATE: 20-MAR-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A
REGISTRATION NUMBER: 32,145
REFERENCE/DOCKET NUMBER: 1107.05064

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 97430 BMB UT
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 486 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5951775e
US-08-821-355A-8

Query Match 1.9%; Score 7; DB 2; Length 486;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
DB 8 GGGAGGG 14

RESULT 58
US-09-003-687A-8
Sequence 8, Application US/09003687A
Patent No. 5998600
GENERAL INFORMATION:

APPLICANT: Barker, Nick
APPLICANT: Clevers, Hans
APPLICANT: Korinek, Vladimir
APPLICANT: Morin, Patrice
APPLICANT: Kinzler, Kenneth
APPLICANT: Vogelstein, Bert
APPLICANT: Sparks, Andrew
TITLE OF INVENTION: Beta Catenin, TCF-4, and APC
TITLE OF INVENTION: Interact to Prevent Cancer
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Ltd.
STREET: 1001 G Street, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20001

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/003,687A
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/821,355
FILING DATE: 20-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Kegan, Sarah A
REGISTRATION NUMBER: 32,145
REFERENCE/DOCKET NUMBER: 1107.05064
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 97430 BMB UT

INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 486 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5998600e
US-09-003-687A-8

Query Match 1.9%; Score 7; DB 2; Length 486;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
DB 8 GGGAGGG 14

RESULT 59
US-09-136-605-8
Sequence 8, Application US/09136605A
Patent No. 6140052
GENERAL INFORMATION:
APPLICANT: He, Tong-Chuan
APPLICANT: Kinzler, Kenneth
APPLICANT: Vogelstein, Bert
TITLE OF INVENTION: Beta Catenin, TCF-4, and APC Interact to
TITLE OF INVENTION: Prevent Cancer
FILE REFERENCE: 1107.75741
CURRENT APPLICATION NUMBER: US/09/136,605A
CURRENT FILING DATE: 1998-08-20
EARLIER APPLICATION NUMBER: 08/821,355
EARLIER FILING DATE: 1997-03-20
EARLIER APPLICATION NUMBER: 09/003,687
EARLIER FILING DATE: 1998-01-06
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 8
LENGTH: 486
TYPE: PRT
ORGANISM: Homo sapiens
US-09-136-605-8

Query Match 1.9%; Score 7; DB 3; Length 486;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
DB 8 GGGAGGG 14

RESULT 60
US-08-821-355A-9
Sequence 9, Application US/08821355A
Patent No. 5851775
GENERAL INFORMATION:
APPLICANT: Barker, Nick
APPLICANT: Clevers, Hans
APPLICANT: Korinek, Vladimir
APPLICANT: Morin, Patrice
APPLICANT: Kinzler, Kenneth
APPLICANT: Vogelstein, Bert
APPLICANT: Sparks, Andrew
TITLE OF INVENTION: Beta Catenin, TCF-4, and APC
TITLE OF INVENTION: Interact to Prevent Cancer
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Ltd.
STREET: 1001 G Street, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20001

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/821,355A
FILING DATE: 20-MAR-1997

```
;
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A
; REGISTRATION NUMBER: 32,145
; REFERENCE/DOCKET NUMBER: 1107.05064
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9299
; TELEFAX: 202-508-9299
; TELEX: 97430 BMB UT
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 511 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 5851775e
; US-08-821-355A-9

Query Match 1.9%; Score 7; DB 2; Length 511;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 GGGAGGG 241
Db 8 GGGAGGG 14

RESULT 61
US-09-003-687A-9
; Sequence 9, Application US/09003687A
; Patent No. 5998600
; GENERAL INFORMATION:
; APPLICANT: Barker, Nick
; APPLICANT: Clevers, Hans
; APPLICANT: Korinek, Vladimir
; APPLICANT: Morin, Patrice
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Vogelstein, Bert
; APPLICANT: Sparks, Andrew
; TITLE OF INVENTION: Beta Catenin, TCF-4, and APC
; TITLE OF INVENTION: Interact to Prevent Cancer
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 1001 G Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fast-SEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/003,687A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/821,355
; FILING DATE: 20-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A
; REGISTRATION NUMBER: 32,145
; REFERENCE/DOCKET NUMBER: 1107.05064
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; TELEX: 97430 BMB UT
; INFORMATION FOR SEQ ID NO: 9;
```

```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 511 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 5998600e
; US-09-003-687A-9

Query Match 1.9%; Score 7; DB 2; Length 511;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 GGGAGGG 241
Db 8 GGGAGGG 14

RESULT 62
US-09-136-605-9
; Sequence 9, Application US/09136605A
; Patent No. 6140052
; GENERAL INFORMATION:
; APPLICANT: He, Tong-Chuan
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Vogelstein, Bert
; TITLE OF INVENTION: Beta Catenin, TCF-4, and APC Interact to
; TITLE OF INVENTION: Prevent Cancer
; FILE REFERENCE: 1107.75741
; CURRENT APPLICATION NUMBER: US/09/136,605A
; CURRENT FILING DATE: 1998-08-20
; EARLIER APPLICATION NUMBER: 08/821,355
; EARLIER FILING DATE: 1997-03-20
; EARLIER APPLICATION NUMBER: 09/003,687
; EARLIER FILING DATE: 1998-01-06
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-136-605-9

Query Match 1.9%; Score 7; DB 3; Length 511;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 GGGAGGG 241
Db 8 GGGAGGG 14

RESULT 63
US-09-252-991A-18266
; Sequence 18266, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: ABRUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18266
; LENGTH: 521
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-18266
```

Query Match 1.9%; Score 7; DB 4; Length 521;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GGGAGGG 241
 Db 152 GGGAGGG 158
 |||||

RESULT 64
 US-08-687-590-27
 ; Sequence 27, Application US/08687590
 ; Patent No. 6255070
 ; GENERAL INFORMATION:
 ; APPLICANT: Willison, Keith Robert
 ; APPLICANT: Kubota, Hiroshi
 ; APPLICANT: Ashworth, Alan
 ; TITLE OF INVENTION: Folding Proteins
 ; NUMBER OF SEQUENCES: 74
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: United States of America
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/687,590
 ; FILING DATE: 31-JUL-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/GB95/00192
 ; FILING DATE: 31-JAN-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9401791.0
 ; FILING DATE: 31-JAN-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9418234.2
 ; FILING DATE: 09-SEP-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Bastian, Kevin L.
 ; REGISTRATION NUMBER: 34,774
 ; REFERENCE/DOCKET NUMBER: 084619-0000000US
 ; INFORMATION FOR SEQ ID NO: 27:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 539 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-687-590-27

Query Match 1.9%; Score 7; DB 3; Length 539;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 269 MLVELSK 275
 Db 90 MLVELSK 96
 |||||

RESULT 65
 US-09-702-705-326
 ; Sequence 326, Application US/09702705
 ; Patent No. 6504010
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Lodes, Michael A.

; APPLICANT: Fanger, Gary
 ; APPLICANT: Vedvick, Tom
 ; APPLICANT: Carter, Darriack
 ; APPLICANT: Retter, Marc
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Fan, Liqun
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; FILE REFERENCE: 210121.478C14
 ; CURRENT APPLICATION NUMBER: US/09/702,705
 ; CURRENT FILING DATE: 2000-10-30
 ; NUMBER OF SEQ ID NOS: 1833
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 326
 ; LENGTH: 539
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-702-705-326

Query Match 1.9%; Score 7; DB 4; Length 539;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 269 MLVELSK 275
 Db 90 MLVELSK 96
 |||||

RESULT 66
 US-09-736-457-326
 ; Sequence 326, Application US/09736457
 ; Patent No. 6509448
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Lodes, Michael A.
 ; APPLICANT: Fanger, Gary
 ; APPLICANT: Vedvick, Tom
 ; APPLICANT: Carter, Darriack
 ; APPLICANT: Retter, Marc
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Wang, Aijun
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; FILE REFERENCE: 210121.478C15
 ; CURRENT APPLICATION NUMBER: US/09/736,457
 ; CURRENT FILING DATE: 2000-12-13
 ; NUMBER OF SEQ ID NOS: 1864
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 326
 ; LENGTH: 539
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-736-457-326

Query Match 1.9%; Score 7; DB 4; Length 539;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 269 MLVELSK 275
 Db 90 MLVELSK 96
 |||||

RESULT 67
 US-09-067-351-1
 ; Sequence 1, Application US/09067351
 ; Patent No. 5994081
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Corley, Neil C.

APPLICANT: Baughn, Mariah
TITLE OF INVENTION: HUMAN KERATINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/067,351
FILING DATE: Herewith
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0511 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 546 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PANCUTUT02
CLONE: 1467090
US-09-067-351-1

Query Match 1.9% Score 7; DB 2; Length 546;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 GGGAGGG 241
Db 108 GGGAGGG 114

RESULT 68
US-09-360-490-1
Sequence 1, Application US/09360490
Patent No. 6221843
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Baughn, Mariah
TITLE OF INVENTION: HUMAN KERATINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/360,490
FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/067,351
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0511 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 546 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PANCUTUT02
CLONE: 1467090
US-09-360-490-1

Query Match 1.9% Score 7; DB 3; Length 546;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 GGGAGGG 241
Db 108 GGGAGGG 114

RESULT 69
US-09-413-814-89
Sequence 89, Application US/09413814
Patent No. 6225064
GENERAL INFORMATION:
APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
APPLICANT: Bristol-Myers Squibb, Co.
APPLICANT: Bayer, Stefan
APPLICANT: Bloecker, Helmut
APPLICANT: Brandt, Petra
APPLICANT: Cino, Paul M
APPLICANT: Dougherty, Brian A
APPLICANT: Goldberg, Steven L
APPLICANT: Hofle, Gerhard
APPLICANT: Mueller, Joachim
APPLICANT: Reichenbach, Hans
TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
TITLE OF INVENTION: heteropolyketide compounds
FILE REFERENCE: PCT/US 99/23535
CURRENT APPLICATION NUMBER: US/09/413,814
CURRENT FILING DATE: 1999-10-07
EARLIER APPLICATION NUMBER: DE 198 46 493.2
EARLIER FILING DATE: 1998-10-09
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 89
LENGTH: 590
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-413-814-89

Query Match 1.9% Score 7; DB 3; Length 590;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 79 TSGSTGT 85
Db 241 TSGSTGT 247

RESULT 70
US-09-413-814-76
Sequence 76, Application US/09413814

```
/ Patent No. 6225064
/ GENERAL INFORMATION:
/ APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
/ APPLICANT: Bristol-Myers Squibb, Co.
/ APPLICANT: Beyer, Stefan
/ APPLICANT: Bloecker, Helmut
/ APPLICANT: Brandt, Petra
/ APPLICANT: Cino, Paul M
/ APPLICANT: Dougherty, Brian A
/ APPLICANT: Goldberg, Steven L
/ APPLICANT: Hofle, Gerhard
/ APPLICANT: Mueller, Joachim
/ APPLICANT: Reichenbach, Hans
/ TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
/ TITLE OF INVENTION: heteropolyketide compounds
/ FILE REFERENCE: PCT/US 99/23535
/ CURRENT APPLICATION NUMBER: US/09/413,814
/ CURRENT FILING DATE: 1999-10-07
/ EARLIER APPLICATION NUMBER: DE 198 46 493.2
/ EARLIER FILING DATE: 1998-10-09
/ NUMBER OF SEQ ID NOS: 107
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 76
/ LENGTH: 591
/ TYPE: PRT
/ ORGANISM: Sorangium cellulosum
US-09-413-814-76

Query Match          1.9%; Score 7; DB 3; Length 591;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 TSGSTGT 85
      |||||
Db 242 TSGSTGT 248

RESULT 71
US-09-252-991A-20134
/ Sequence 20134, Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ CURRENT FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 20134
/ LENGTH: 609
/ TYPE: PRT
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20134

Query Match          1.9%; Score 7; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 283 LIENHD 289
      |||||
Db 157 LIENHD 163

RESULT 72
US-08-894-454-163
/ Sequence 163, Application US/08894454
/ Patent No. 6544784
/ GENERAL INFORMATION:
/ APPLICANT: VAN DEN VEN, W.J.M.
/ APPLICANT: SCHOENMAKERS, H.F.P.M.
/ TITLE OF INVENTION: MULTIPLE-TUMOR ABERRANT GROWTH
/ TITLE OF INVENTION: GENES
/ NUMBER OF SEQUENCES: 164
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: The Webb Law Firm
/ STREET: 700 Koppers Building, 436 Seventh Avenue
/ CITY: Pittsburgh
/ STATE: PA
/ COUNTRY: USA
/ ZIP: 15219-1818
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/894,454
/ FILING DATE: 15-AUG-1997
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/EP/00716
/ FILING DATE: 19-FEB-1996
/ APPLICATION NUMBER: 95200390.3
/ FILING DATE: 17-FEB-1995
/ APPLICATION NUMBER: 95201951.1
/ FILING DATE: 14-JUL-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Johnson, Barbara E
/ REGISTRATION NUMBER: 31,198
/ REFERENCE/DOCKET NUMBER: 702-971100
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 412-471-8815
/ TELEFAX: 412-471-4094
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 163:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 612 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
US-08-894-454-163

Query Match          1.9%; Score 7; DB 4; Length 612;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 GGEGDFL 307
      |||||
Db 62 GGEGDFL 68

RESULT 73
US-09-328-352-4474
/ Sequence 4474, Application US/09328352
/ Patent No. 6562958
/ GENERAL INFORMATION:
/ APPLICANT: Gary L. Breton et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
/ TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: GTC99-03PA
/ CURRENT APPLICATION NUMBER: US/09/328,352
/ CURRENT FILING DATE: 1999-06-04
/ NUMBER OF SEQ ID NOS: 8252
/ SEQ ID NO 4474
/ LENGTH: 615
/ TYPE: PRT
/ ORGANISM: Acinetobacter baumannii
US-09-328-352-4474

Query Match          1.9%; Score 7; DB 4; Length 615;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
```

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 79 TSGSTGT 85
|||||
Db 164 TSGSTGT 170

RESULT 74

US-09-345-473E-46
; Sequence 46, Application US/09345473E
; Patent No. 6558903
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin
; TITLE OF INVENTION: NO. 6558903el Kinases and Uses Thereof
; FILE REFERENCE: 35800/183781
; CURRENT APPLICATION NUMBER: US/09/345,473E
; CURRENT FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-345-473E-46

Query Match 1.9%; Score 7; DB 4; Length 627;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 236 GGAGGGP 242
|||||
Db 22 GGAGGGP 28

RESULT 75

US-09-345-473E-47
; Sequence 47, Application US/09345473E
; Patent No. 6558903
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin
; TITLE OF INVENTION: NO. 6558903el Kinases and Uses Thereof
; FILE REFERENCE: 35800/183781
; CURRENT APPLICATION NUMBER: US/09/345,473E
; CURRENT FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-345-473E-47

Query Match 1.9%; Score 7; DB 4; Length 627;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 236 GGAGGGP 242
|||||
Db 22 GGAGGGP 28

Search completed: December 17, 2003, 06:36:30
Job time : 14.8538 secs

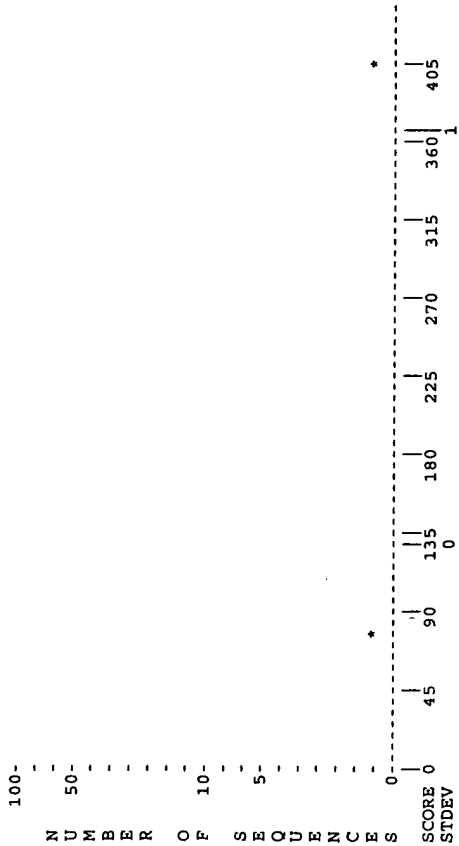
> O <
O| O IntelliGenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-09-805-550-2.res made by jdelaval on Wed 17 Dec 103 6:22:09-PST.

Query sequence being compared: US-09-805-550-2 (1-405)
Number of sequences searched: 2
Number of scores above cutoff: 2

Results of the initial comparison of US-09-805-550-2 (1-405) with:
File : US09805550.pep



PARAMETERS

Similarity matrix Unitary
Mismatch penalty 1
Gap penalty 1.00
Gap size penalty 0.05
Cutoff score 0
Randomization group 0

K-tuple 2
Joining penalty 20
Window size 32

SEARCH STATISTICS

Scores: Mean 240 Median 77 Standard Deviation 232.64
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00

Number of residues: 773
Number of sequences searched: 2
Number of scores above cutoff: 2

The scores below are sorted by initial score.
Significance is calculated based on initial score.
A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Length	Score	Init. Opt.	Sig. Frame
US-09-805-550-2	Sequence 2, Application US	405	405	0.71	0

1. US-09-805-550-2 Sequence 2, Application US 405 405 0.71 0
The list of other best scores is:

Sequence Name	Description	Length	Score	Init. Opt.	Sig. Frame
US-09-805-550-4	Sequence 4, Application US	368	76	125	-0.70 0

1. US-09-805-550-2 (1-405)
US-09-805-550-2 Sequence 2, Application US/09805550

Sequence 2, Application US/09805550
GENERAL INFORMATION:
APPLICANT: Mahajan, Pramod B.
TITLE OF INVENTION: Rad23 Genes and Uses Thereof
FILE REFERENCE: 0964D
CURRENT APPLICATION NUMBER: US/09/805,550
CURRENT FILING DATE: 2001-03-13
PRIOR APPLICATION NUMBER: 09/413,574
PRIOR FILING DATE: 1999-10-06
PRIOR APPLICATION NUMBER: 60/109,728
PRIOR FILING DATE: 1998-11-23
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 405
TYPE: PRT
ORGANISM: Zea mays

Initial Score = 405 Optimized Score = 405 Significance = 0.71
Residue Identity = 100% Matches = 405 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
MCLNVTLLKGTNFEIEASPDASVADVKKRIETTCQSTYRADQQMLIYQGLKDKDETTLESNGVAENSLVI
MCLNVTLLKGTNFEIEASPDASVADVKKRIETTCQSTYRADQQMLIYQGLKDKDETTLESNGVAENSLVI
X 10 20 30 40 50 60 70
80 90 100 110 120 130 140
MLSKAKASSSGASTATTAKAPATLAQPAAPVAPASVARTPTQAPVATETAPPVQPOQAAATVAAATDDA
MLSKAKASSSGASTATTAKAPATLAQPAAPVAPASVARTPTQAPVATETAPPVQPOQAAATVAAATDDA
80 90 100 110 120 130 140
150 160 170 180 190 200 210
DVYQAASNLVFGNNLEQTIQIILDMGGTWERDVTVRALRAAYNNPERAIDYLSGIPENVEAQPVARAPA
DVYQAASNLVFGNNLEQTIQIILDMGGTWERDVTVRALRAAYNNPERAIDYLSGIPENVEAQPVARAPA
150 160 170 180 190 200 210
220 230 240 250 260 270 280
AGQNTNQAASPAQPAVALPVQSPASAGPNANPLNLPFGVPSGSGNPGVVPAGSGALDALROLPOQAL
AGQNTNQAASPAQPAVALPVQSPASAGPNANPLNLPFGVPSGSGNPGVVPAGSGALDALROLPOQAL
220 230 240 250 260 270 280
290 300 310 320 330 340 350 360
LQLVQANPQIILQPMLELQELGKONPQILRIQENQAEFLRLVNESPEGGPGNLTGLAAAVPQTLTTPERE
LQLVQANPQIILQPMLELQELGKONPQILRIQENQAEFLRLVNESPEGGPGNLTGLAAAVPQTLTTPERE
290 300 310 320 330 340 350 360
370 380 390 400 X
AIIQRLGGMGNRELVLVEFFACNCKDEELTANYLLDHGHEFDQQQ
AIIQRLGGMGNRELVLVEFFACNCKDEELTANYLLDHGHEFDQQQ
370 380 390 400 X

> O <
O| O IntelliGenetics
> O <

GENALIGN - Multiple Sequence Alignment Program
Release 5.4

Wed 17 Dec 103 6:20:37-PST

Solution Parameters:

Amino Alphabet = Identity
Output line length = 80
Compress = Off
Histogram = Off
Randomization = Off

AMINO-Res-length = 2
Deletion-weight = 5.00
Length-factor = 0
Matching-weight = 1.00
NUCLEIC-Res-length = 4
Spread-factor = 50

Clustered order of selected sequences:

1. US-09-805-550-2 (1-405)
2. US-09-805-550-4 (1-368)

Region Alignment: (listed in Clustered order)

```
US-09-805- 1 MKLnVKTlKGtNFfEasPdasvadVKrIEtTQqstYradQQmLiYqGKiLKDEtTLES
US-09-805- 1 MKLcVKTlKGtNFfEirvPndtImaVKnIEiQKdsYpwgQQLlInGKvLKDEsTlEe
consensus MKL-VKTlKGT-FEI---P-----VK--IE--QG---Y---QQ-LI--GK-LKDE-TLE-

US-09-805- 62 NgVaEneFLVlMSKakassSGaStattakaplaqPAapvAPaasvartPtQAPvatae
US-09-805- 62 NkVnEdgFLVvWMLSK gktSG ST gssssghentPAtzqAP pleapQAP qppv
consensus N-V-E--FLV-MLSkaK---SGaStat-----PA-----APaas----P-QAPV----

US-09-805- 123 tappsvQPqaPaatvaatddadvysqAASNLvFGnNleqtIqildMGGGtWerDtVvRA
US-09-805- 114 apittsQegilPA gapnthdnAASNLlEGrNvdtIInQlmeWGGGsWdkDKvQRA
consensus -----QP---Paatvaat-----AASNL--G-N-----I-Q---MGGG-W--D-V-RA

US-09-805- 184 LRAAYNNPERAidLYSGIPenveaqPvarPaasGQtnnqqaasPQapavalPvqpsPaSa
US-09-805- 169 LRAAYNNPERAveLYSGIP vtaeiavPiGQ gAnttdraPtgeaglsG
consensus LRAAYNNPERA--LYSGIPenve-----P--GQtnnqqaas-A-----P-----S-

US-09-805- 245 gPnanPLnLFPQGvpgsNpvyvPgAGsGaldALrglPQFQAillqlvqanPQILQPMlQe
US-09-805- 218 iPNtApLdLFPQG aSN aggGAGgGpLDfLRnnPQFQAvremVhtNPQILQPMlVe
consensus -PN--PL-LFPQGvpsg-SNPg---GAG-G-LD-LR--PQFQA----V--NPQILQPMl-E

US-09-805- 306 lGKQNPQILRLlIQnqaEFLvNesPegGpGgnilglaaavPqtltVTPEEeEAIqRLS
US-09-805- 273 LsKQNPQILRLlIeNhdeFLqlLNEpfEGGgGdldqpeedemPhaisVTPEEeEAIqRLS
consensus L-KQNPQILRLI-EN--EFL-L-NE--EGG-G-----P-----VTPEE-EAI-RLE
```

```
US-09-805- 367 gMGFnRelVlEvFfACnkdbEELtANYLLdHghsfDdqqg
US-09-805- 334 eMGFdRarViEaFlACdrnEELaANYLLLeHagReD
consensus -MGF-R--V-E-P-AC---EEL-ANYLL-H--E-Ddqqg
```

Alignment score = 7.00

Scoring matrix:

	1	2
1		-1
2		
